

IJCSIS Vol. 10 No. 3, March 2012
ISSN 1947-5500

International Journal of Computer Science & Information Security

© IJCSIS PUBLICATION 2012

Editorial

Message from Managing Editor

International Journal of Computer Science and Information Security (IJCSIS) provide a forum for publishing empirical results relevant to both researchers and practitioners, and also promotes the publication of industry-relevant research, to address the significant gap between research and practice.

Being a fully open access scholarly journal, original research works and review articles are published in all areas of the computer science including emerging topics like cloud computing, software development etc. It continues promote insight and understanding of the state of the art and trends in technology. To a large extent, the credit for high quality, visibility and recognition of the journal goes to the editorial board and the technical review committee.

Authors are solicited to contribute to the journal by submitting articles that illustrate research results, projects, surveying works and industrial experiences. The topics covered by this journal are diversified. (See monthly Call for Papers)

For complete details about IJCSIS archives publications, abstracting/indexing, editorial board and other important information, please refer to IJCSIS homepage. IJCSIS appreciates all the insights and advice from authors/readers and reviewers. Indexed by the following International Agencies and institutions: EI, Scopus, DBLP, DOI, ProQuest, ISI Thomson Reuters. Average acceptance for the period January-March 2012 is 31%.

We look forward to receive your valuable papers. If you have further questions please do not hesitate to contact us at ijcsiseditor@gmail.com. Our team is committed to provide a quick and supportive service throughout the publication process.

A complete list of journals can be found at:

<http://sites.google.com/site/ijcsis/>

IJCSIS Vol. 10, No. 3, March 2012 Edition

ISSN 1947-5500 © IJCSIS, USA & UK.

Journal Indexed by (among others):



IJCSIS EDITORIAL BOARD

Dr. Yong Li

School of Electronic and Information Engineering, Beijing Jiaotong University,
P. R. China

Prof. Hamid Reza Naji

Department of Computer Engineering, Shahid Beheshti University, Tehran, Iran

Dr. Sanjay Jasola

Professor and Dean, School of Information and Communication Technology,
Gautam Buddha University

Dr Riktesh Srivastava

Assistant Professor, Information Systems, Skyline University College, University
City of Sharjah, Sharjah, PO 1797, UAE

Dr. Siddhivinayak Kulkarni

University of Ballarat, Ballarat, Victoria, Australia

Professor (Dr) Mokhtar Beldjehem

Sainte-Anne University, Halifax, NS, Canada

Dr. Alex Pappachen James (Research Fellow)

Queensland Micro-nanotechnology center, Griffith University, Australia

Dr. T. C. Manjunath

HKBK College of Engg., Bangalore, India.

Prof. Elboukhari Mohamed

Department of Computer Science,
University Mohammed First, Oujda, Morocco

TABLE OF CONTENTS

1. Paper 29021266: Integrating LBS, GIS and SMS Technologies for the Effective Monitoring of Road Network (pp. 1-6)

S. El houssaini, A. Badri

Laboratoire d'Electronique, Electrotechnique, Automatique & Traitement de l'Information

Faculté des Sciences et Techniques de Mohammedia, Université Hassan II Mohammedia-Casablanca, B.P.146, Mohammedia, Morocco

2. Paper 27111119: Improving Information Security in E-Banking by Using Biometric Fingerprint (pp. 7-12)

Mahmoud Mohammed Mahmoud Musleh, Karama M.A. Nofal, Ismail Idrissa Ba, Jamaludin Ibrahim

Department of Information Systems, International Islamic University Malaysia, IIUM, Kuala Lumpur, Malaysia

3. Paper 30041180: Mobile WiFi-Based Indoor Positioning System (pp. 13-22)

Mike Ng Ah Ngan, Mohammed Abdul Karim, Behrang Parhizkar, Arash Habibi Lashkari

Faculty of Information & Communication Technology, LIMKOKWING University, Cyberjaya, Selangor, Malaysia

4. Paper 27021212: Kekre's Wavelet Transform for Image Fusion and Comparison with Other Pixel Based Image Fusion Techniques (pp. 23-31)

Dr. H. B. Kekre, MPSTME, SVKM'S, NMIMS University

Dr. Tanuja Sarode, Computer Engineering Department, Thadomal Shahani Engineering College

Rachana Dhannawat, Computer Sci. & engg. Department, S.N.D.T. University, Mumbai

5. Paper 27021213: A Survey on Building Intrusion Detection System Using Data Mining Framework (pp. 32-36)

V. Jaiganesh² and M. Thenmozhi³, Dr. P. Sumathi¹,

² Assistant Professor, Department of Computer Science, Dr.N.G.P. Arts and Science College, Coimbatore

³ Assistant Professor, Department of Information Technology, Faculty of Engineering, Avinashilingam University for Women, Coimbatore

¹ Assistant Professor, Department of Computer Science, Chikkanna Government Arts College, Tirupur

6. Paper 27021218: Trusted On-demand Distance Vector Routing for Ad hoc Networks (pp. 37-44)

Md. Humayun Kabir, Bimal Kumar Pramanik, Somlal Das, Subrata Pramanik and Md. Ekramul Hamid

Department of Computer Science and Engineering, University of Rajshahi, Rajshahi-6205, Bangladesh

7. Paper 29021231: An Intrusion Detection System Framework for Ad Hoc Networks (pp. 45-48)

Arjun Singh, Dept. of Computer Science & Engineering, Sir Padampat Singhania University, Udaipur, India

Surbhi Chauhan, Dept. of Computer Science & Engineering, Amity University, Noida, India

Kamal Kant, Dept. of Computer Science & Engineering, Amity University, Noida, India

Reshma Doknaia, Sr. Software Engineer, BMC Pvt. Ltd., Pune, India

8. Paper 29021234: Data Mining Techniques: A Key for detection of Financial Statement Fraud (pp. 49-57)

Rajan Gupta¹ and Nasib Singh Gill²

¹ Research Scholar, Dept. of Computer Sc. & Applications, Maharshi Dayanand University, Rohtak (Haryana) - India.

² Head, Dept. of Computer Sc. & Applications, Maharshi Dayanand University, Rohtak (Haryana), India.

9. Paper 29021240: Performance Comparison of Assorted Color Spaces for Multilevel Block Truncation Coding based Face Recognition (pp. 58-63)

Dr. H.B. Kekre, Senior Professor, Computer Engineering Department, MPSTME, SVKM's NMIMS, (Deemed-to-be University), Mumbai, India

Dr. Sudeep Thepade, Associate Professor, Computer Engineering Department, MPSTME, SVKM's NMIMS, (Deemed-to-be University), Mumbai, India

Karan Dhamejani, Sanchit Khandelwal, Adnan Azmi

B.Tech Students, Computer Engineering Department, MPSTME, SVKM's NMIMS, (Deemed-to-be University), Mumbai, India

10. Paper 29021243: Step Tapered Waveguide with Cylindrical Waveguide (pp. 64-66)

Harshukumar Khare, M.E (EXTC) Final year student, TEC, Nerul, Navi-Mumbai

Prof. R. D. Patane, Asst. Proffessor (EXTC), TEC, Nerul, Navi-Mumbai

11. Paper 29021248: A Panoramic Approach on Software Quality Assurance Proposed By CMM and XP (pp. 67-71)

*CH. V. Phani Krishna^{*1}, Dr. G. Rama Krishna^{*2} and Dr. K. Rajasekhara Rao^{*3}*

¹ Associate professor, CSE Department, KL University, Guntur dt., India.

² Professor, CSE Department, KL University, Guntur Dt., India

³ Dean of student and faculty welfare, KL University, Guntur Dt., India.

12. Paper 29021254: Developing Multi-Platform Package for Remote System Administration (pp. 72-76)

Rawaa Putros Polos Qasha, Department of Computers Sciences, College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq

13. Paper 29021256: An Efficient Automatic Attendance System Using Fingerprint Reconstruction Technique (pp. 77-82)

Josphineleela. R¹, Ramakrishnan. M²

¹ Research Scholar, Sathyabama University, Chennai,

² Department Of Information Technology, Velammal Engineering College, Chennai

14. Paper 29021260: Towards More Realistic Mobility Model in Vehicular Ad Hoc Network (pp. 83-90)

Dhananjay S. Gaikwad, Mahesh Lagad, Prashant Suryawanshi, Vaibhav Maske

Computer Engineering Department, HSBPVT'S GOI, Parikrama College of Engineering, Kashti. India- 414701

15. Paper 29021263: Image Classification in Transform Domain (pp. 91-97)

Dr. H. B. Kekre, Professor, Computer Engineering, Mukesh Patel School of Technology Management and Engineering, NMIMS University, Vileparle(w) Mumbai 400-056, India

Dr. Tanuja K. Sarode, Associate Professor, Computer Engineering, Thadomal Shahani Engineering College, Bandra(W), Mumbai 400-050, India

Jagruti K. Save, Ph.D. Scholar, MPSTME, NMIMS University, Associate Professor, Fr. C. Rodrigues College of Engineering, Bandra(W), Mumbai 400-050, India

16. Paper 29021264: Analysis of Stock Marketing with SOAP service using Python (pp. 98-102)

P. Asha, Research Scholar, Computer Science and Engineering Department, Sathyabama University, Chennai, Tamilnadu, India.

Dr. T. Jebarajan, Principal, Kings College of Engineering, Chennai, Tamilnadu, India.

Kathiresan, Technical Lead Consultant, Motorola Solutions, Bangalore, Karnataka, India.

17. Paper 20021208: Accurate Face Recognition Using PCA and LDA (pp. 103-112)

*Sukhvinder Singh *, Mtech CSE (4th sem), Sri Sai College Of Engg. & Tech., Pathankot*

Meenakshi Sharma, HOD CSE, Sri Sai College Of Engg. & Tech., ,Pathankot

Dr. N Suresh Rao, HOD CSE, Sri Sai College Of Engg. & Tech., Jammu University

18. Paper 20021209: Robust & Accurate Face Recognition using Histograms (pp.113-122)

Sarbjeet Singh ¹, Meenakshi Sharma ², Dr. N.Suresh Rao ³
SSCET Pathankot^{1,2}, Jammu University³

19. Paper 31011219: X.509 Authentication Services to Enhance the Data Security in Cloud Computing (pp. 123-125)

Surbhi Chauhan, Amity University, Noida, India

Kamal Kant, Amity University, Noida, India

Arjun Singh, Sir Padampat Singhania University, Udaipur, India

20. Paper 27021215: An Intelligent Spam-Scammer Filter Mechanism Using Bayesian Techniques (pp. 126-139)

Olushola D. Adeniji, Olubukola Adigun and Omowumi O. Adeyemo

^{1,2&3} Department of Computer Science, University of Ibadan, Ibadan, Nigeria

21. Paper 27021216: New Weather Forecasting Technique using ANFIS with Modified Levenberg-Marquardt Algorithm for Learning (pp. 140-147)

I.Kadar Shereef, Sree Saraswathi Thyagaraja College, Pollachi. Coimbatore, Tamil Nadu, India.

Dr. S. Santhosh Baboo, Dwaraka Doss Goverdhan Doss Vaishnav College, Chennai, Tamil Nadu. India

Integrating LBS, GIS and SMS Technologies for the Effective Monitoring of Road Network

Souad El houssaini*, Abdelmajid Badri

Laboratoire d'Electronique, Electrotechnique, Automatique & Traitement de l'Information
Faculté des Sciences et Techniques de Mohammedia, Université Hassan II Mohammedia-Casablanca, B.P.146
Mohammedia, Morocco
souad_elhoussaini@yahoo.fr ; abdelmajid_badri@yahoo.fr

Abstract—This paper presents an integrated framework of Geographic Information System (GIS), Android Platform and a Relational Database Management System (RDBMS) equipped with interactive communication capabilities. The model integrates the design of the database and the management of implementation of the monitoring system which includes the operations of query and analysis using the web and desktop applications. This study aims to apply techniques of analysis of the road network in a GIS to collect geographic data on the monitoring station and the roads. The information on road infrastructure is not only useful for locating monitoring stations, but it is also important to guide a station to follow the shortest path to achieve the objectives of management and routing. Optimal routes based on the minimum cost are identified using Dijkstra's algorithm. This paper also presents a software development on Android Platform which applies Cell Identifier method for improving the accuracy of location, it is not necessary to have an Internet connection as the requesting emergency can use a Short Message Service (SMS) to request an urgent service. The proposed system should be an effective and intelligent tool for a rapid intervention and to improve the monitoring of the road network which can eventually be extended to a national infrastructure of GIS. Simulated test cases have been carried out for network of Mohammedia City in Morocco.

Keywords: *GIS; Location; Routing; SMS; Android; Cell Identifier.*

I. INTRODUCTION

Road safety has long been a major concern in the road sector. Road accidents can cause serious injury or death; these effects can also lead to significant economic losses for the payee. The accelerated rapid development of wireless network and mobile computing technologies has increased the convenience of mobile information services for solving real-life problems, such as monitoring of road accidents. In general, the Location-Based Service (LBS), a software application which retrieves information about where a mobile device is located, uses GPS (Global Positioning System) to indicate the geographical position of the mobile device. Since all mobile devices cannot be equipped with GPS receiver, the

GPS may not be suitable for location of the requesting emergency. In order to solve this problem, Cell Identifier which indicates mobile device position by using station base information is introduced in this research. With the facilities of Android that provides LBS components for retrieving information about where a mobile device is located [1], a system that retrieves the location of the mobile was developed. The results from this process are composed of mobile's particulars information: MCC (Mobile Country Code), MNC (Mobile Network Code), Cell ID (Cell Identifier) and LAC (Location Area Code), this information is sent by SMS from the mobile of requesting emergency to the monitoring system. Teleoperator can use these clues to locate the phone and inform the monitoring stations in order to save property and people.

This study aims also to implement and evaluate a methodology based on GIS (Geographic Information System) to determine optimal routes of the road network using key information items based on cost of distance. With this paper we will try to help and fill that gap, presenting a decision tool for monitoring stations for location and routing. This approach saves time. This article is comprised of the following parts: Part 2, which introduces available tools; Part 3, which describes the structure of the proposed system; Part 4 which examines the implementation; Part 5, which discusses the experiment result; and finally Part 6, the conclusion.

II. AVAILABLE TOOLS

A. Global Positioning System (GPS): Outdoor Localization System

GPS is a system used for determining the position of interesting objects such as person, pets or vehicles. This system receives satellite signals and calculates the position of mobile device of which a SIM card is installed for sending the co-ordinate (latitude and longitude) of its position to the recipient [1].

While GPS is widely used in outdoor localization, it does not perform effectively in indoor localization. This is because it lacks the ability to pierce through building wall and requires custom infrastructures for every area in which localization is to be performed [2, 3].

B. Global System for Mobile Communication (GSM): Indoor Localization System

GSM is a digital mobile telephony system which is widely used in most part of the world. GSM identifies mobile device position by using Cell Identifier method which retrieves information from cell tower. The Cell Identifier method provides many benefits for localization, utilizing the widely and most accessible network infrastructure in most parts of the world [3]. Therefore this method is applied in our system for finding the position of requesting emergency.

C. Android Platform

Android is a platform for mobile device developed by Google. It provides a complete set of software development: operating system, tools and APIs necessary to begin developing applications [4]. The java-based programming makes Android widely used in developing mobile application [1]. In this research, Android Cell ID API was applied to obtain the Cell ID of an Android mobile.

D. Geographic Information System (GIS)

Geography information systems have been improving since 1970s. GIS is an essential tool for location mapping, dynamic condition visualization, and decision making [5–7]. Geospatial data are useful in monitoring response to accidents. The analysis of real-time data could be achieved through GIS during the response phase to support visualization and automation for efficient decision making. Research has been conducted in GIS that focused on areas such as shortest path analysis [8, 9]. This shows the great potential of GIS applications to facilitate the possibility of having a response time shorter if the geospatial information is implemented in the initial phase of response to accidents.

III. STRUCTURE OF THE PROPOSED SYSTEM

A. Requirements

Figure 1 depicts a use case and the use case diagram for the Management System for Road Safety (MSRS). Use cases are used for documentation of functional requirements and for communication between stakeholders and developers. This is a common practice in software engineering that ensures the software developers understand the requirements. Thus, the developed system is expected to address the requirements set by the stakeholders.

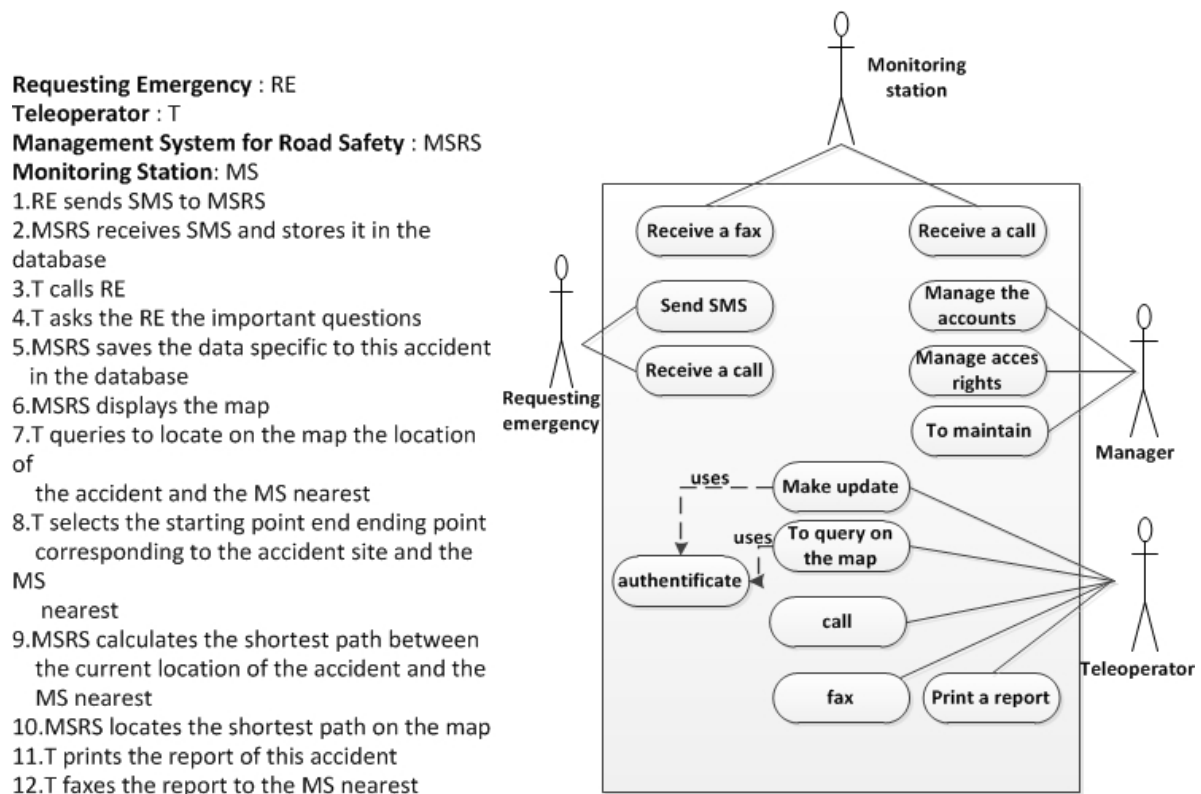


Figure 1: Use case diagram and use case for MSRS.

B. System design

The web framework based on a three-tier architecture consisting of the client layer, middleware layer and the layer of the database (Figure 2). These components together provide a unified interface for consultation data, request and decision making for users, the database is accessed through the Internet, in such a way that the user does not need to be aware of the location of the database, it is sufficient that the user is able to consult, add and modify the data as needed.

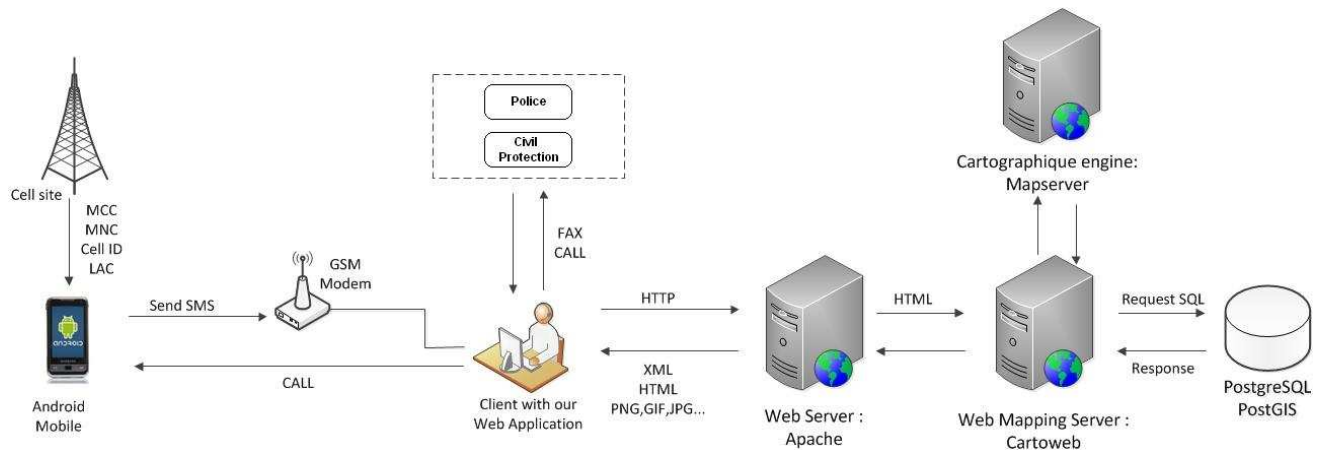


Figure 2: The structure of the proposed system.

First of all, after installing our program in Android mobile, the requesting emergency clicks the button "send sms". Moreover, the position of the mobile will be sent in real-time to monitoring system by retrieving mobile information from the closest cell site consisting of MCC, MNC, Cell ID and LAC, which will be further used for retrieving latitude and longitude of the mobile. GSM modem receives this SMS, the monitoring system saves in the database the data contained in this SMS. A GSM modem can be an external modem device, a PC card installed in a notebook computer, or a standard GSM mobile phone, in our case we used GSM mobile phone to test our application. The Android service is applied to send all of information to monitoring system. The advantage of this system is to provide clues of location for teleoperator to notify the monitoring station to serve the requesting emergency.

Application uses PostgreSQL as the database management system with the geospatial extension PostGIS. Additional to relational queries, PostGIS provides spatial queries to the users. The GIS products used are Mapserver and Cartoweb, Mapserver is used here just as library PhpMapScript, Cartoweb is a solution designed for the web, it allows its architecture CartoClient / CartoServer to answer

several web service in the world of free Web Mapping. It fits easily with Apache and php5. Web-based GIS users can use a Web browser to navigate maps and to complete basic spatial analysis. The requests from the user are sent to clients by way of HTML forms. The form is passed to the Web server Apache and a gateway at the Web server passes the request to GIS server Cartoweb, then Cartoweb queries the database.

IV. IMPLEMENTATION

A. Implementation of Dijkstra's Algorithm

Dijkstra's algorithm calculates the least accumulated cost between the destination node and every other node, and then finds the least-cost path from any origin nodes to the destination node.

The Dijkstra's algorithm is very similar to the A* algorithm. The cost function (c) used to evaluate shortest paths in the Dijkstra algorithm is augmented by an estimator function that is used to estimate the shortest path between two given graph nodes [i.e., $c(s, d) = g(s, v) + h(v, d)$, where $g(s, v)$ is the cost from source s to v and $h(v, d)$ is the heuristic estimated cost from v to the destination d]. The estimator function is a heuristic function that can be chosen arbitrarily. If the estimator function is 0, A* turns into Dijkstra's algorithm [10].

In our system, the routing service has been implemented using the Dijkstra's algorithm in the road network of the city of Mohammedia. The algorithm was implemented with PHP5 in Cartoweb environment. The version of Cartoweb used is 3.5.0 and runs with Windows XP operating system. Cartoweb

not only allows the handling of road maps online using user-friendly interfaces, but it also allows to implement the routing plugin nome "pgrouting", it runs client side and server side. The user can define the points of beginning and end of the shortest path by selecting their names (names of streets). Whenever a routing operation is performed, their names are passed as parameters from the client interface to the server. The server connects to the database, it prepares the request, it computes the shortest path connecting the given points, draw it on the map with a different color. Figure 3 depicts the shortest path between Boulevard "11 Janvier" and Boulevard "Sebta" in the road network in Mohammedia.

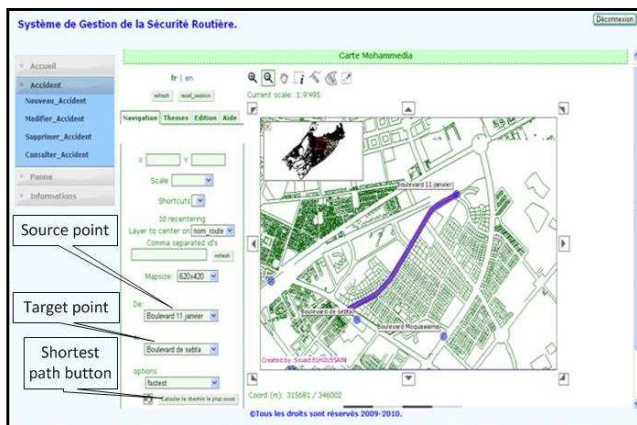


Figure 3: Visualization of the shortest path.

B. Receiving SMS Through a Computer

To send and receive SMS, a GSM modem [11] of high band rate will be needed so that a large number of messages could be receive at high speed at every moment. In order to get connected to the GSM modem through a computer, the standardized AT commands must be used. The set of commands used for controlling modems is called AT command. Every AT command includes a result code which specifies the status of the command and a reply containing the data returned by the modem. AT commands usually begin with the prefix "AT" [12].

Figure 4 shows how a short message is received via our application developed in VB6 and saved in the database.

The application to receive SMS and to save it in the database was developed by two different methods: one using AT commands and the other with the android platform if the GSM modem does not support AT commands.

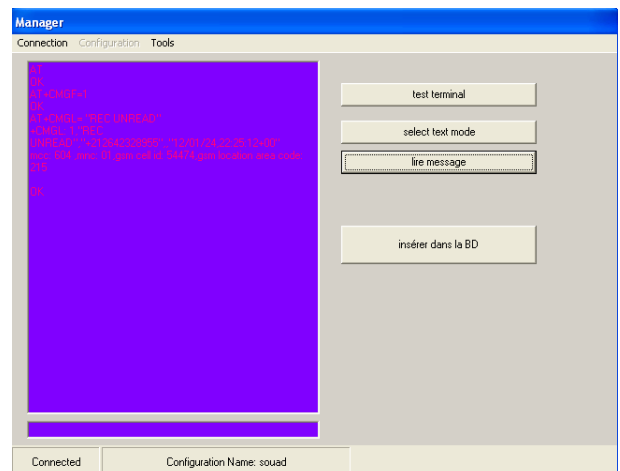


Figure 4: Receiving a short message through a computer using AT Command.

C. Other Implementations of Decision Model

The choice was preferentially oriented means "Open-source" such as Mapserver, Cartoweb, PostgreSQL and Android. To implement the application, we used an Object-Oriented (OO) methodology (Unified Modeling Language—UML). The developed web application is organized around a main window, with all the functionalities accessible in this window, through the toolbar, or the menu in a way easily understandable by users.

The tool developed is composed of a set of Graphical User Interfaces (GUI). It was implemented for Windows platforms and has an open architecture which allows an easy integration of new functionality. The teleoperator uses GUI to make the interpretation of information easier (Figure 5).

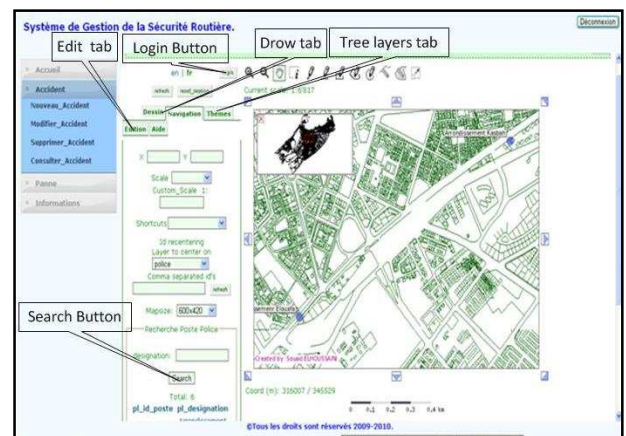


Figure 5: The GUI of different visualizations tabs.

The move towards Android technology is rapidly approaching. With the facilities of Android that provides LBS components for retrieving information about where a mobile device is located, a system that retrieves the location of mobile was developed. In order to locate mobile device, Cell Identifier of GSM network is applied. This step was developed by applying Android Cell ID API which can obtain the Cell ID of an Android device. The result of our system consists of mobile's identification information: MNC, MCC, Cell ID and LAC (Figure 6).



Figure 6: Information retrieved from the mobile of requesting emergency.

V. EXPERIMENT RESULT

Observing the results given in table 1 shown that after testing the mobile location system in several different places around Mohammedia (Figure 7) by comparing our measurements obtained from a mobile equipped with our application and those of GPS, the comparison shows that the system can retrieve the requesting emergency mobile information and sent it to the monitoring system at 99.9% accuracy, one can conclude that the proposed system can be successfully applied in real application for monitoring system of road accidents.

VI. CONCLUSION

In this paper, we described an intelligent system offering a solution to the treatment of emergency accidents in the city of Mohammedia in Morocco for automatic monitoring. An operation of great significance for this treatment is the delivery the monitoring station to the sites asking for help to save property and people. The system has been tested in a real case study, its architecture involves an integrated framework of Geographic Information System (GIS), Android Platform and a Relational Database Management System (RDBMS) equipped with interactive communication. The objective of this research is to provide a system for location and routing. The improved efficiency by GIS reduces the task of maintaining paper maps. The Web-based GIS framework facilitates the orientation of the monitoring station to the location of the accident. The Android platform applies

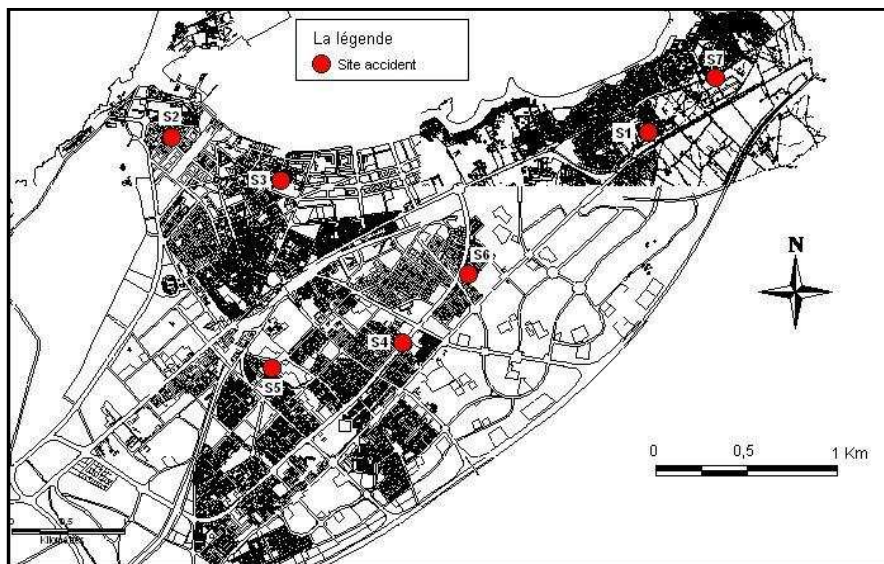


Figure 7: Geographic location of supposed sites of road accidents.

TABLE 1: System Testing Result

Supposed sites of road accident	Latitude Application	Longitude Application	Latitude GPS	Longitude GPS
S1	33°42'35''N	007°21'19''O	33°42'40''N	007°21'19''O
S2	33°42'30''N	007°24'02''O	33°42'36''N	007°24'03''O
S3	33°42'20''N	007°23'24''O	33°42'24''N	007°23'25''O
S4	33°41'45''N	007°22'41''O	33°41'38''N	007°22'42''O
S5	33°41'33''N	007°23'32''O	33°41'30''N	007°23'27''O
S6	33°41'54''N	007°22'18''O	33°41'58''N	007°22'20''O
S7	33°42'53''N	007°21'09''O	33°42'56''N	007°20'56''O

the Cell Identifier to improve the efficiency of localization. There are several advantages of the developed system. First, time saving and flexibility are important merits of the system. The system is Object-Oriented (OO) and understandable, it has potential to be integrated with the other roads networks and to be expanded to a national base, so the model can be extended to all cities of Morocco using the technologies: GIS, Android and RDBMSs. As for future work, an algorithm can be developed to calculate the minimum distance between the location of the road accident and monitoring stations neighbors to determine the monitoring station nearest.

ACKNOWLEDGMENTS

This work falls within the scope of telecommunication projects. We would like to thank the Department of technology of the MESFCRST for financing our projects.

REFERENCES

- [1] S. Sukaphat. "An Implementation of Location-Based Service System with Cell Identifier for Detecting Lost Mobile", *Procedia Computer Science*, vol. 3, pp. 949–953, 2011.
- [2] GPS Beginner's Guide, Resource. [Online]. Available : http://www8.garmin.com/manuals/GPSGuideforBeginners_Manual.pdf. [Last Accessed: February 15, 2012].
- [3] B. Rao and L. Minakakis. "Evolution of mobile location-based services", *Communication of the ACM*, vol. 46, pp. 61-65, 2003.
- [4] S. Sukaphat. "Creating of Mobile Search System for Traffic Inquiry", in *Proc. 10th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD 2009)*, 2009, pp. 417 - 420.
- [5] M. Kataoka. "GIS for Homeland Security", ESRI Press, Redlands, CA, 2007.
- [6] A.R. Pradhan, D.F. Laefer and W.J. Rasdorf. "Infrastructure management information system framework requirements for disasters, ASCE", *Journal of Computing in Civil Engineering*, vol. 21, pp. 90–101, 2007.
- [7] M.K. Lindell, and C.S. Prater. "A hurricane evacuation management decision support system (EMDSS)", *Natural Hazards*, vol. 40, pp. 627–634, 2007.
- [8] M.P. Kwan, and J. Lee. "Emergency response after 9/11: the potential of real-time 3d GIS for quick emergency response in micro-spatial environments", *Computing Environment and Urban Systems*, vol. 29, pp. 93–113, 2005.
- [9] J. Han, Z. Yong and K.W. Dai. "The approach for shortest paths in fire succor based on component GIS technology", *SPIE: The International Society for Optical Engineering, Geoinformatics 2007, Geospatial Information Technology and Applications*, vol. 6754, 2007.
- [10] I. Maglogiannisa and S. Hadjiefthymiades. "EmerLoc: Location-based services for emergency medical incidents", *International journal of medical informatics*, vol. 76, pp. 747–759, 2007.
- [11] S. Voglioukas and M. Roumeliotis. "A System for Basic-level Network Fault Management Based on the GSM Short Message Service (SMS)", *IEEE International Conference on Communications*, 2001.
- [12] F. Givehki and A. Nicknafs. "Mobile control and management of computer networks using SMS services", *Telematics and Informatics*, vol. 27, pp. 341–349, 2010.
- [13] Android Developers. [Online]. Available : <http://developer.android.com/index.html>. [Last Accessed: January 31, 2012].

Improving Information Security in E-Banking by Using Biometric Fingerprint

A CASE OF MAJOR BANK IN MALAYSIA

Mahmoud Mohammed Mahmoud Musleh

Department of Information Systems
International Islamic University Malaysia, IIUM
Kuala Lumpur, Malaysia
mahmd.musleh@gmail.com

Ismail Idrissa Ba

Department of Information Systems
International Islamic University Malaysia, IIUM
Kuala Lumpur, Malaysia
ba1isma1a@yahoo.fr

Karama M.A. Nofal

Department of Information Systems
International Islamic University Malaysia, IIUM
Kuala Lumpur, Malaysia
karama.nofal@gmail.com

Jamaludin Ibrahim

Department of Information Systems
International Islamic University Malaysia, IIUM
Kuala Lumpur, Malaysia
jamaludinibrahim@kict.iium.edu.my

Abstract— In this paper biometric fingerprint technology will define and discuss as new best approach identification and authentication customers for online internet banking, and how biometric fingerprint will improve the internet banking protect its assets. Background will be produced to present how authentication and identification have developed and improved through the applications successful that have implemented biometric technology to protect its asset; then a case of major bank in Malaysia will be taken as a case study. By answering the question, why does biometric fingerprint need to come forefront as a great method of authentication in online banking environment? The findings have found that there are reasons and factors for higher security as a near perfect and biometric fingerprint authentication will be indicated to be the solution to answer this call.

Keywords- Biometric Fingerprint; E-banking; Information Security; Online Banking; Biometric Technology

I. INTRODUCTION

Millions of dollars are being invested in the developed of e-banking systems worldwide, and it is of paramount importance that these systems are fully utilized by potential customers. However, there remains reluctance by consumers to accept e-banking because of the perceived risk security financial and time. Therefore, banks need to better understand their customers and respond to developments in internet technology in a way that incorporates their customers' requirements and addresses their concerns [16].

There are three major types of authentication commonly used; the first type is something that you know such as a password, PIN or a piece of personal information.

Secondly, it is something that you have such as a smart card or token. The last type is something that you are such as a biometric [10]. On the other hand, many organizations are using the internet as a new distribution channel to provide their customers a good service such as internet banking [14]. This channel needs to be secure and trusted not only to protect the customer information from fishing or hacking, but also provide data integrity; and to ensure providing the services in a safety way. Therefore, Information security has become a major concern for banks to conserve their customers' assets. In addition, everyday there are updates of security to face the challenges that have faced internet banking; in parallel, there are intruders who think every moment to attack others. This paper will focus on biometric fingerprint technology as a solution to deter the threats that concerns e-banking security as much as possible.

A. Background of Study

Issues with biometric device include accuracy and failure. Some researchers mentioned that biometric still have negative impact denying access to unauthorized user. What happen if the user is wearing a bandage on the finger of authentication? For this scenario some device provide password. One of the issues regarding biometric is cost effective, in today organization user work in the office, at home, and in hotel, airport, and internet café. If you decide to purchase biometric device for all employees, how many device will you buy? [2]. While others have seen a biometric fingerprint is a powerful way of deciding who can gain access to our most valuable system in this modern world; despite biometric fingerprint are successfully adopted in areas such as Automatic Teller Machine (ATM) [4, 15]; however, there is a lack of implementation to online banking environment [17].

According to online banking of two major banks in Malaysia, customers used username and password to access their accounts. However, the difference is that one uses TAC number to authenticate when the customer needs to make transaction, and another thing the customer has to answer questions that he knew the words exactly when customer subscribed in the internet banking and he put his own answer.

As a result, biometric fingerprint as a near perfect security is still in its infancy for most major banks in Malaysia. Those did not take a risk in order to achieve biometric solution to enhance their security systems. Some opponents argue that password only authenticate a password but not the user. Password can be forgotten and forged by the hackers. Password does not provide a non-repudiation security service which means to ensure that transferred message has been sent and received by the parties claiming to have sent and received the message and also password is very vulnerable [3]. Biometric method will basically authenticate the person and internet banking that must have a non-repudiation security service to ensure that customer cannot later deny his transaction. Some security expert argue that biometric is the only true user authentication because of it physical authentication [2]. As some people will see, biometric will not be the best choice for every one [5]. On the other hand, biometric technology appeals to many banking organizations as a near perfect solution to such security threats [17]. Therefore, the biometric fingerprint technology is the best method to protect and secure online banking assets. The banks should adopt biometric fingerprint technology as a near perfect solution to such security threats of internet banking in particular for major bank in Malaysia.

B. Scope of Study

This study will focus on factors that influence the bank to be ready to use biometric fingerprint to authenticate the user when make transaction on internet banking. Existing literatures will be used to design the study. Although there are many researches talk about biometric technology available in various literatures, but this study will focus on only the biometric fingerprint to investigate whether the major bank in Malaysia ready to use biometric fingerprint in internet banking. Qualitative will be used to carry this paper, sample will be chosen, and afterwards gathering information and analysis will be performed.

C. The Significance

The paradigm shift from something that the users know to something that the users are; online banking requires the development and implementation of trustworthy security procedure [7]. Therefore, the newly emerged service such as fingerprint biometric to use it in the internet banking for authentication and identification and rapidly increasing penetration rates of internet banking to be as near perfect security are the motivators of this study [13].

- Biometric fingerprint considers as a new technology in online banking environment which means it needs a lot of efforts and resources to be used.

- The biometric fingerprint has become a significant phenomenon in recent times, it has various advantages and benefits in both organization and customer.

II. LITERATURE REVIEW

A. An overview of Information Security

With the rapid growth of Information and Communication Technology (ICT), information security becomes more pervasive in everyday lives while there are many channels and methods to attack of websites with this great development of information security. One of the threats to web authentication is phishing, where a phishing attack is a type of social engineering attack, designing users' authentication credentials by spoofing the login page of a trusted web site [9]. However, some banks in Malaysia use TAC number by sending it to customer's mobile to authenticate the user when he make transaction and others use the questions that the customer has already known his own answer when he subscribed in internet banking.

Some opponents argue that the information which the person knew such as a password only authenticate a password but not the user and can be forgotten and forged. The information such as the question that supposed the customers knew can be forgotten and forged by the hackers [2]. Furthermore, Password does not provide a non-repudiation security service and the passwords are easily broken with the programs that available on the internet that help to break the password and may be people will choose easily remembered and easily gassed password such as name of their relative, date of birth or phone number [12, 3].

B. Online Banking Security with Biometric

Online banking demands the development and implementation of trustworthy security procedure [7]. This requirement needs to design effective method that works efficiently via which users or customers can be verified and authenticated in a remote environment.

The biometric fingerprint has become an important phenomenon in recent times, it has various advantages and benefits in both organization and customer [13]. However, it is yet to be adopted by major bank in Malaysia.

Many studies have been conducted on biometric fingerprint technology, and many researchers have discussed the influence that biometric technology as a perfect solution for many purposes [4, 5, 13, & 17]. In contrast, there is still a lack of research on the factors or the ability of banks to be ready to use biometric fingerprint in internet banking to authenticate the user.

C. Definitions of Terms Used

1) Information Security in Business

In business information security helps managers to govern, monitor and secure the information from malware changes and removals or unauthorized access. The main aims of Information security in business is to protect the confidentiality from a competitor or media and integrity that is to ensure that

the information is not changed or modified as well to ensure the availability of the information when needed or in an event of a disaster [12]. Many businesses are merely depending on information deposited in computers; personal information, and details that may all be warehoused on a database. Without this information, it would often be very hard for a business to function. Information security systems need to be implemented to protect this asset [2].

Nowadays, there are many types of threats available on the internet that need to be enforced to ensure business goals. Based on Proctor, 2002 organizations and their information system and network are faced with security threat from wide range of sources including computer fraud, espionage, sabotage, and vandalism [18]. Cause of damage such as code, computer hacker and denial for services attach have become more common and increasingly spreading in the World Wide Web.

2) Security Policy

A policy is a document that summarizes rules that must be abided by the organization. Security policy is the backbone of the security architecture without a policy you cannot protect your information [2]. In addition, policies allow the organization to reduce cost and eliminate accountability. Written policy works as the means of communicating company guidelines to the customer [11]. Furthermore, policy defines how security should be implemented, this comprise proper configuration. Thus policy provides the rules that govern how system should be configured and how customers of an organization should act in normal circumstance and react during unusual situation. Some examples recommended for biometric Policy; do not share your fingerprint device with any person, any obvious act of fraud or guessing the fingerprint the services will be terminated report to the bank immediately when the device is stolen.

3) Biometric Fingerprint

The term biometrics is used to describe physical dimensions and/or behaviour characteristics which are essential and unique to the human being; and it can be utilized to verify the identity of a person. These characteristics include fingerprint, hand geometry, facial characteristics, iris, retina, personal scent and DNA, while behaviour features include handwriting, keystroke, voice and gait. Physiological characteristics can be measured and recognized [8]. Biometric fingerprint technology is considered one of the most secure and convenient authentication tool. It cannot be stolen, borrowed, or forgotten, and forged [10].

III. THE DIFFICULTIES AND CHALLENGES THE PROJECT FACED

Getting information from the banks is very challenging because of the sensitivity asset; in addition, the bank policy stated that it is illegal to reveal the customers' information and

their strategy plan. Furthermore, the bank delayed to respond our request to meet the human resource manager. Since there is a high competition among the banks, so every bank wants to keep their strategies from the researchers and press. Asset protection is the biggest challenge in information security systems of the banks. They have sensitive information such as customers' information and their credit card details, which need to be secured. Therefore, protecting information against leakage has become more complex and difficult when an opponent who is authorized to view the data or information about the processes of the security system [1].

Based on Harris & Spence, 2002 banks are increasingly threatened by the leakage of sensitive information which can be available to impostors or competitors. Furthermore, Banks want to ensure that information assets such as the security system, trade secrets, software code, designs, architectures, and algorithms are not leaked and abused [6]. Also, they want protection against leakage of internal confidential information, which can damage the customers' trust to the company brand.

According to these reasons the major bank in Malaysia rejected to give us any information about their e-banking security system; to avoid leakage of information which can compromise their security system and affect their competitiveness of protecting the confidential information of their customers.

IV. PROPOSED E-BANKING SECURITY SYSTEM PROCESSES

A. Authentication processes to access the account

The diagram shows that the authentication process consists of two stages. First of all, the user needs to verify his/her username and password, if the username and the password are accepted; the browser will direct the user to the second stage of authentication but if the username and the password are not accepted the browser will ask the user to reinsert valid username and password.

Secondly, this stage is the most significant one which is the authentication stage by using the biometric fingerprint technology. The user needs to verify his/her fingerprint by using fingerprint reader which is connected to his/her own personal computer (Figure 1). The fingerprint server will match the user fingerprint with the bank's fingerprints database; if it is accepted the browser will direct the user to access his/her accounts.

These two stages of authentication protect the customer information from unauthorized reading that means confidentiality of the customer which is very important from the customer's perspective because it saves him/her from failing under the threat of the malicious people.

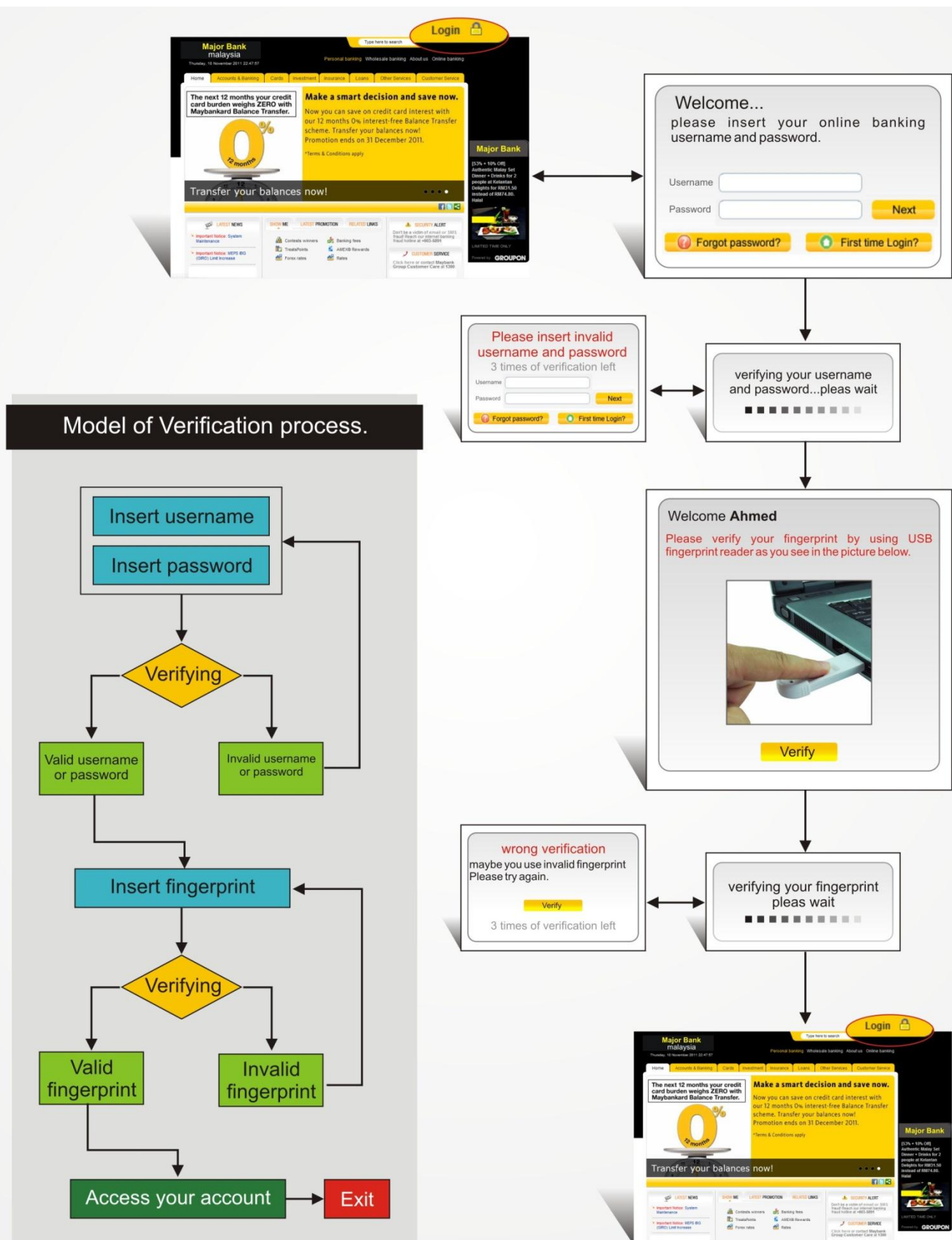


Figure 1. Authentication processes to access the account

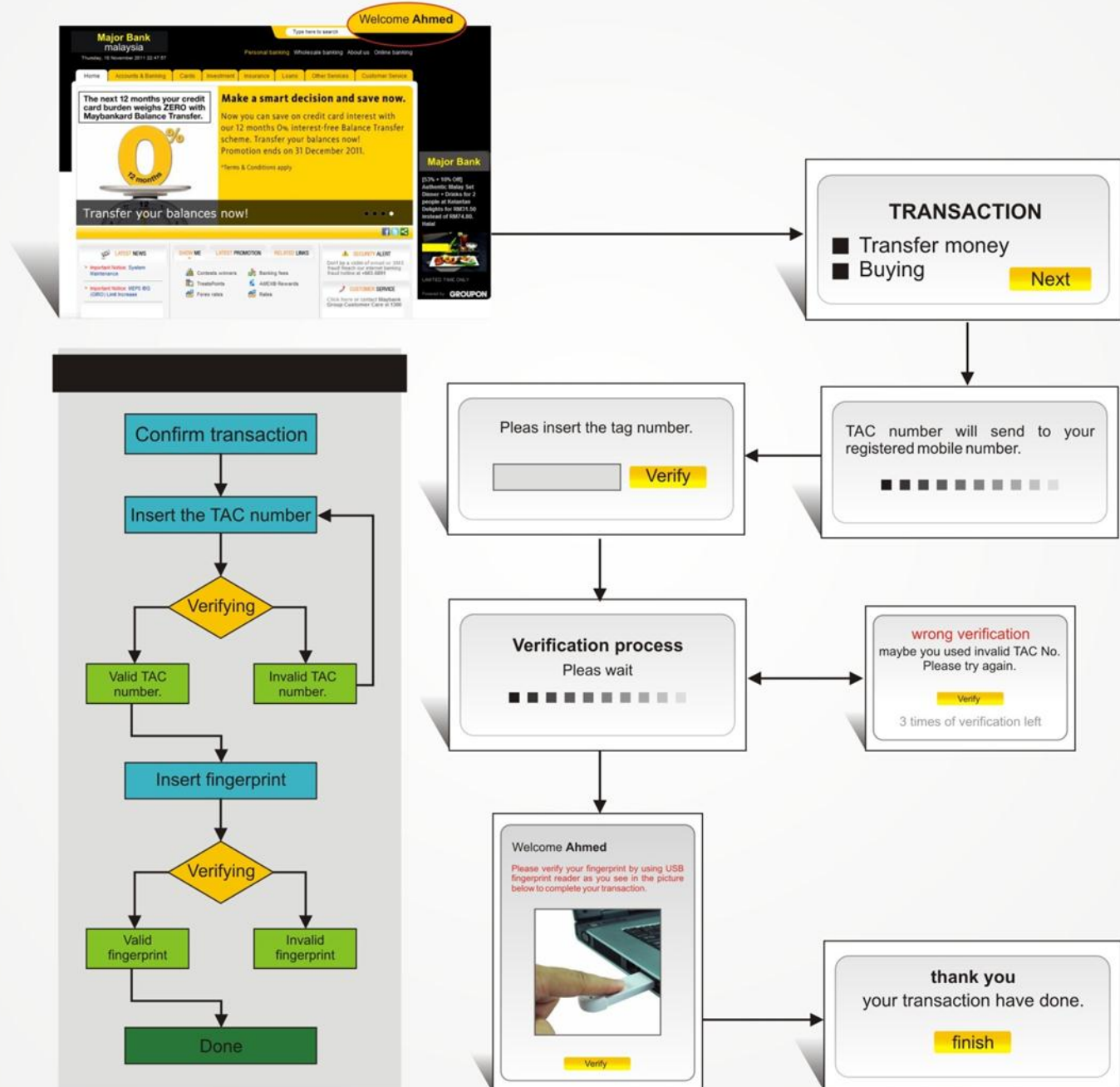


Figure 2. Authentication processes of transaction

B. Authentication processes of transaction

This process consists of two stages of authentication that the customer needs to confirm his/her transaction first stage is by using TAC, second one is by using biometric fingerprint technology (Figure 2).

Authentication process by using Transaction Authorization Code (TAC), e-banking system will send TAC automatically to the customer's mobile number, which is registered in the database of the bank system. The customer will receive text message (SMS) includes on Transaction Authorization Code (TAC). Therefore, after inserting the TAC the system will verify it, if it is accepted the browser will direct the customer to confirm his/her fingerprint again to complete the transaction.

The confirmation processes of transaction should be very secure because it protected the customer account from unauthorized changing, editing, or writing. This process is called integrity which is required to protect the customer assets.

CONCLUSION

Information security is becoming ubiquitous whether is logical or physical. It is an essential approach for every organization to protect its asset from intruders and malware. Most of the banks experienced many threats and abuse in their system. Information security ensures the confidentiality of information. The numbers of users of online banking has significantly increased; therefore, biometric fingerprint will be used to enforce the authentication and identification of the user with username and password as an approach. Researchers argued that biometric fingerprint is secure mechanism used to authenticate the person because password only verifies the username but not the physical identity such as person fingerprint. In addition, customers, employees are the weakest layer in information security.

As a result, policies will be utilized on how configure the device as well as training the people about awareness of security. The purpose of policy is to protect not only the company asset from threats whether internal or external but also to reduce cost and eliminate legal liability to employees. This paper will give the researchers the insight about biometric as the powerful tool and perfect solution for authentication.

REFERENCES

- [1] M. I. Abbadi & M. Alawneh. "Preventing Insider Information Leakage for Enterprises", The Second International Conference on Emerging Security Information, Systems and Technologies, IEEE journal, pp. 99-160, DOI: 10.1109/SECURWARE.2008.14, 2008.
- [2] A. Andress. *Surviving security*, 2004.
- [3] R. Ayoub & C. Rodriguez. "A Best Practices Guide to Fingerprint Biometrics: Ensuring a Successful Biometrics Implementation", White paper, 2011. Retrieved Nov., 2011 from: <http://www.frost.com/prod/servlet/cpo/240303611>
- [4] S. Debbarma & S. Das. "Designing a Biometric Strategy (Fingerprint) Measure for Enhancing ATM Security in Indian E-Banking System". *IJICT Journal*, Volume 1 No. 5, pp. 197-203, 2011.
- [5] A. J. Harris & D. C. Yen. "Biometric authentication: Assuring access to information", *Information Journal of Management and Computer Security*, Emerald Group Publishing Limited, 10(1), 12-19, 2002.
- [6] L. Harris & L. J. Spence. "The ethics of e-banking". *Journal of Electronic Commerce Research*. VOL. 3, NO. 2, 2002.
- [7] D. Hutchinson & M. Warren. "Security for Internet banking: a framework", *Logistics Information Management*, Emerald Group Publishing Limited, 16(1), pp.64 – 73, 2003.
- [8] P. Jones. "Biometrics in retailing", *International Journal of Retail & Distribution Management*, Vol. 35 No. 3, 2007 pp. 217-222, 2007.
- [9] C. K. Karlof. "Human Factors in Web Authentication", University of California, Berkeley, Technical Report No. UCB/EECS-2009-26, 2009. Retrieved Oct., 2011 From: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-26.pdf>
- [10] S. Liu & M. Liu. "A Practical Guide to Biometric Security Technology", *IEEE Journal*, 3(1), PP. 23-32, 2001.
- [11] E. Maiwald. *Fundamentals of Network Security*, 2004.
- [12] M. Merkow & J. Breithaupt. *Information Security: Principles and Practices*, 2006.
- [13] J. E. Mills, M. Meyers, & S. Byun. "Embracing broad scale applications of biometric technologies in hospitality and tourism: Is the business ready?", *Journal of Hospitality and Tourism Technology*, Emerald Group Publishing Limited, Vol. 1 No. 3, pp. 245-256, 2010.
- [14] M. Nami. "E-Banking: Issues and Challenges", *ACIS International Conference on Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing*, 2009.
- [15] N. C. Sickler & S. J. Elliott. "An evaluation of fingerprint image quality across an elderly population vis-a-vis an 18-25 year old population", *IEEE*, PP. 68-73, 2005.
- [16] R. Tassabehji & M. A. Kamala. "Improving E-Banking Security with Biometrics: Modelling user attitudes and acceptance", *IEEE Journal*, pp. 1-6, DOI: 10.1109/NTMS.2009.5384806, 2009.
- [17] S. Venkatraman & I. Delpachitra. "Biometric in banking security: A case study", *Information Journal of Management and Computer Security*, Emerald Group Publishing Limited, 16(4), 415-430, 2008.
- [18] P. E. Proctor. *The Secured Enterprise Protecting your Information Asset*, 2002.

Mobile WiFi-Based Indoor Positioning System

MIKE NG AH NGAN

Faculty of Information & Communication Technology,
LIMKOKWING University
Cyberjaya, Selangor, Malaysia
mike.ng@limkokwing.edu.my

MOHAMMED ABDUL KARIM

Faculty of Information & Communication Technology,
LIMKOKWING University
Cyberjaya, Selangor, Malaysia
abdulkarim@limkokwing.edu.my

BEHRANG PARHIZKAR

Faculty of Information & Communication Technology,
LIMKOKWING University
Cyberjaya, Selangor, Malaysia
hani.pk@limkokwing.edu.my

ARASH HABIBI LASHKARI

Faculty of Information & Communication Technology,
LIMKOKWING University
Cyberjaya, Selangor, Malaysia
a_habibi_L@hotmail.com

Abstract—Navigation system wherever built inside a GPS device or on a mobile phone has been proved to be very useful for outdoor environment. The device gives you your exact position and shows you the direction to your destination. But nowadays, it is clearly seen that a navigation system may be beneficial for indoor environment as well. This paper illustrates a mobile application which will be able to estimate the position of a user within a building by using WiFi technology.

Keywords—WiFi, WiFi positioning System, Indoor Positioning System

1. INTRODUCTION

Navigation system wherever built inside a GPS device or a mobile phone has been proved to be very useful for outdoor environment. The device gives you your exact position and shows you the direction to your destination. But nowadays, it is clearly seen that a navigation system may be beneficial for indoor environment as well.

The design of GPS was based partly on similar ground-based radio navigation systems, such as LORAN and the Decca Navigator developed in the early 1940s, and used during World War II. In 1956 Friedwardt Winterberg proposed a test of general relativity using accurate atomic clocks placed in orbit in artificial satellites. To achieve accuracy requirements, GPS uses principles of general relativity to correct the satellites' atomic clocks. [1]

The first satellite navigation system, Transit, used by the United States Navy, was first successfully tested in 1960. It used a constellation of five satellites and could provide a navigational fix approximately once per hour.

Global Positioning system was created and realized by the U.S. Department of Defense (DOD) and was originally run with 24 satellites. It was established in 1973 to overcome the limitations of previous navigation systems and is the most prominent contribution in determining position of user and in routing him to his destination. This system uses satellites to triangulate the location of the GPS device. Though this system

has made a good impression in terms of accuracy and is the preferred location based system for outdoor positioning, when it comes to indoor environment, GPS has proved to be inefficient. The reason for its inefficiency is that in order for GPS to perform a triangulation, the device needs to be in line-of-sight from the satellites. Moreover, GPS system has a low

Precision which make it not suitable for indoor areas [2]. Therefore, when it comes to indoor positioning system, other alternatives such as Bluetooth, WiFi, RFID and Infrared Red are more preferable.

The good thing is that all the wireless technologies mentioned above are available on mobile phones. If you look at the mobile phones being unleashed nowadays such as the Nokia N97 or the iPhone from Apple, they both come with built in Bluetooth and WiFi connectivity. These features are indispensable in mobile phones as they helps to send data wirelessly or to connect to the internet wherever you are (taking into consideration that the place has WIFI available).

Among these well-known wireless technologies, the one which mark out from the others is WiFi technology. In most indoor environment such as airports, universities and shopping mall, WIFI is available and is most of the time free. Therefore, anyone using a mobile device with built-in WiFi can connect to the access points and browse the internet easily. Being free and easily accessible is a great advantage as the mobile application should be a low-cost application and accessible to as many user as possible.

The rest of the paper is organized as follows. Section two, explained about the various indoor positioning systems built in the past, what are the advantages and disadvantages of the systems. Section three describes the methodology used for the mobile application. In section four, the implementation of the system is discussed. Section five, shows the testing of the system and as for the conclusion and future works, it is shown in section six.

II. RELATED WORK

A. Different types of wireless technologies used.

The earliest location system was the Active Badge developed at Olivetti Research Laboratory where the user was wearing a badge that emitted infrared signals. [2] Every 10 seconds, a unique identifier is communicated to fixed receivers. The data is then sent to a central server that provides an API. [3] The accuracy of the location depends on the number of receivers. The two limitations of this method are that it requires line-of-sight between the receivers and the badge and infrared has a short-range transmission signal.

RFID has also demonstrated its capability in location-based systems. One famous location sensing system using RFID technology is known as mTag. The mTag architecture uses fixed RFID readers located within the environment and a passive RFID tag attached to a mobile phone or PDA. [4] The disadvantage of using RFID is that the cost of deploying and implementing this kind of system can be very high.

The pervasive adoption of WiFi in indoor environments has provided an opportunity to develop indoor positioning systems that will not require investing in specialized hardware. Some of the well-known systems using WiFi are RADAR, Herecast and PlaceLab.

Radar is one of the first indoor positioning systems based on IEEE 802.11 wireless network. The system, developed by Microsoft research, uses the Radio Frequency Signal strength to measure the distance between the Access Point and the Mobile station. [5] The RADAR system includes two phases, the Training Phase and the Online Phase. In the training phase, an area is divided into a 1x1 meter grid where the signal strength measurements of the access points are taken at each intersection. The mean of the signal strengths which have been obtained, is recorded to create a radio map to be used in the online phase. In the Online phase, when the user looks for its location, the mobile station will detect and record the signal strength from as many access points as possible. Then, the signal strength received will be compared to the radio maps to determine the location of the user.

Herecast is another system using the WLAN technology. [6] It allows the WiFi-enabled client device to determine its location by listening from signals from known access points within the building environment. The system creates a database where the MAC address of the access point is stored together with the symbolic name of the location. In the localization process, the position of the user is the one associated with the access point with the strongest signal strength. The weakness of the system is if an access point is faulty or has been removed, the position of the user may be distorted.

The PlaceLab system is similar to Herecast in that it allows the client device to automatically obtain its location by listening to signal from access points. PlaceLab stores the MAC address broadcast by each access point as well as its longitude and latitude in the client device. Therefore, for when the client device receives a signal from each of the access points, the location is calculated as the average of retrieved longitude and latitude. [7]

Using the 802.11 WiFi signals for location estimation have attracted many researchers as the infrastructure has already been deployed widely in commercial buildings. In addition,

WiFi is appropriate in indoor environment and users are not required to rely on specially developed wireless receivers.

B. Problem with WiFi

Using WiFi technology to estimate the position of the user is one of the most appropriate and profitable because in mostly every public building such as airport, schools and shopping malls, the presence of IEEE 802.11 b/g access points is made available. Therefore, implementing a WiFi-based system would be easier. Unfortunately, though using WiFi for indoor positioning systems has shown promising results, it is not without drawbacks. WiFi signal is a very sensitive signal which can be affected by people, furniture and other architectural components in the indoor environment.

Body Effect

William et al. presents some of the negative effects that decrease the accuracy of using WiFi to estimate the position of the user. The first one is Body effect. WLAN uses 2.4GHz frequency carrier and FCC regulation requires WLAN to operate at low power which is 1 watt or 30 dBm. Since the penetration power is noticeably low, positioning performance can be severely affected. To be more precise, when a user is holding his mobile phone, the path between the PDA and the access point can be obstructed by the user. Therefore, the effect of the human body can make signal strength drop by 10-15 dBm. [8]

In the paper "Properties of Indoor Received Signal Strength for WLAN Location Fingerprint", they have studied the effect of the user's body. They have measured the signal at a specific location which was about 7 m from the access point and was not in line-of-sight for two hours. The first hour, the data were collected with the presence of the user while in the second hour, the user was not present. The histogram below clearly demonstrates the distribution of the RSS with and without the presence of the user. The presence of the user has significantly changed the standard deviation from 0.68 to 3.00 dBm and the mean from -70.4 dBm to -71.6 dBm. [9]

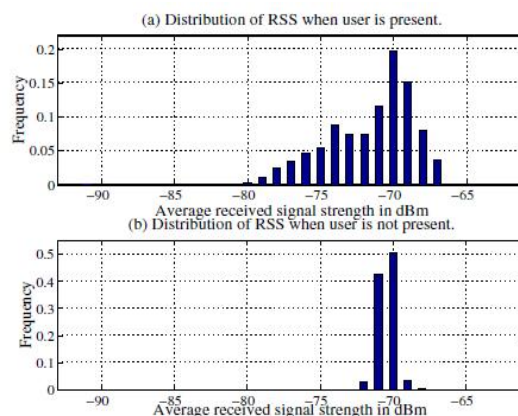


Fig 1: Comparison of histogram of RSS [34]

Trailing Effect

Trailing effect is another issue when dealing with WLAN. Basically, WLAN driver use sliding window to cache beacon messages up to 10 seconds according to the driver design. Consequently, if a user is walking away from an access point, even after he is out of range of the radio signal, the access point will still be visible until timeout occurs. [8]

Signal Aliasing

Signal aliasing refers to two points that are far apart physically but may be close together in signal space. This usually happen because of the complex indoor propagation environment. For instance, the signal strength at a point close to an AP may be similar to another point which is far away essentially because the former point is receiving an obstructed signal due to work while the latter point receives an unobstructed signal. Placement of APs in the building layout is very essential in solving this problem. [10]

C. Techniques for locating mobile station

The idea of locating mobile station was first introduced by Figel et al. in 1969 when they tried to locate a vehicle by using signal attenuation method. Ever since, researches have been done on finding other ways of locating mobile station. Some of these location techniques are received signal strength from Figel et al. in 1969; angle detection by Porter in 1971; and arrival time measurement by Staras and Honicknan in 1977.

Received Signal Strength (RSS)

Signal strength method which is based on signal attenuation is the distance between the Access points and the mobile station. The distance can be calculated either at the Mobile station or the Base station.

Lin et al. in 2004 [11] proposed a mobile location system which used weighted centroid method based on the ratios of distance between the access points and the mobile station derived from the difference of signal attenuation. The benefit of this proposed method is that it does not require perfect path loss and shadowing models. Also, this system can fit directly into the indoor infrastructure without any hardware modification.

P. Bahl and V. Padmanabhan have also developed a system called RADAR which is based on Received Signal Strength. The system collect the RSS from all detected Access points and compared it to the tuple already stored in the Radio Map using search techniques that computes the Euclidean distance between each SS tuple and then choose the one with the minimal distance.

Time of Arrival

Time of arrival (TOA) is referred as a multilateral method that is used to locate the position of a mobile station by measuring the time that it takes for a signal to travel from the mobile station to the base station. Generally, in traditional geometric interpretation, TOAs generate circles whose intersections give the estimate location of the transmitter.

In the paper, "A New Approach to the Geometry of TOA Location" Caffery proposed a new geometrical interpretation in which instead of using circular LOPs to determine the

position of the transmitter, straight lines of position is used. The straight LOPs come from a simple observation regarding the geometry of the system and are not obtained from linearization. [12]

Time Difference of Arrival

Similarly to TOA, time difference of arrival (TDOA) uses the same concept except that it uses time differences measurement rather than absolute time measured. Also, TDOA requires a minimum of three nodes for its most basic operation. The figure below demonstrates a system diagram of how TDOA can be implemented in WLAN. In this system, all the APs need to listen to the same client which is a limitation of this system since APs around a specific client can be set to various frequency channels and therefore can only listen to their selected frequency channel. [13]

D. Localizations' techniques

In this section, 4 types of localization techniques have been discussed.

Weighted Center-of-Gravity Algorithm

The approach in this algorithm is that a value is assigned to every participating access point or node. Given (n) elastic cords connected between the actual position and every access points. So, AP with more tension will attract the approximated position closer to itself. The tension is calculated based on the inverse power law and the power value (α) is variable and approximated in training phase which can be formulated as:

$$\begin{cases} \hat{x} = \frac{x_0/rss_0^\alpha + x_1/rss_1^\alpha + \dots + x_n/rss_n^\alpha}{1/rss_0^\alpha + 1/rss_1^\alpha + \dots + 1/rss_n^\alpha} & (1a) \\ \hat{y} = \frac{y_0/rss_0^\alpha + y_1/rss_1^\alpha + \dots + y_n/rss_n^\alpha}{1/rss_0^\alpha + 1/rss_1^\alpha + \dots + 1/rss_n^\alpha} & (1b) \end{cases}$$

Eq 1: Weighted Center-of-Gravity Algorithm

Where $\hat{y}$ and x represent the estimated coordinates, x_n , y_n and rss_n represents n-th AP position and its corresponding RSS. This method is suitable for mobile devices as the memory footprint is very small since the calculation only requires the location of APs and their environment value (α). [8]

Triangulation

Triangulation is normally used in GPS system. For each AP, a circle is formed with the radius of signal strength and centered at the AP. The circle shows locus where the user may be situated. To estimate position, intersection points are collected and permuted to form triangles where the centroid of the smallest triangle refers to the position of the user. Triangulation does not require high computation but the disadvantage of this method is that if there are too few candidates to form triangles, the result may not be too accurate. Moreover, if the global signal level fluctuates, triangulation may not adapt to the new level.

Smallest M-Vertex Polygon Algorithm (SMP)

Smallest M-Vertex Polygon is a method similar to triangulation in terms of candidates promoted by each neighbour AP but instead of performing space transformation between signal space and world space, it performs estimation with a discrete approach. For instance, instead of finding smallest triangle, SMP estimated the location by determining the smallest polygon. Based on the prepared sample database, each AP endorses a group of candidate location with the same signal strength. Then, assuming M neighbour APs are involved, a M-vertex polygon is constructed from the candidates where each AP will promote a list of candidates for every vertex. Therefore the estimated location is determined as the centroid of the smallest polygon. [8]

Fingerprint Algorithm

The fingerprint approach is based on remembering various radio environs at various locations which are called as markers. These radio environs or snapshots are composed of network address and RSS of nearby APs. In the offline phase, this information together with the marker location is stored in the database. Afterwards, in the positioning phase, the current snapshot taken by the mobile device will be compared to every snapshot in the database. The accuracy of fingerprint highly depends on the separation between markers. Fingerprint-based positioning model has higher precision than propagation model. Moreover, compared to propagation-based model, fingerprint-based model avoids the hard work of finding a general propagation model. The fundamental idea of fingerprint-based system is to look for the nearest neighbour in the signal space by calculating the distance, more precisely, the Euclidean distance between the location of fingerprint already stored in the database with the current RSS tuple obtained at the receiver.

III. METHODOLOGY

Based on the previous section which was the Related Work, various existing systems on indoor positioning have been deeply analyzed to extract the necessary requirements, the strength, the weaknesses and the techniques that are the most appropriate in terms of easy deployment and accuracy. According to the research, a Fingerprint-based system using WiFi technology is more suitable for the proposed system. This segment will explain in detail the methods that will be used for the development of the WiFi-Based Indoor Positioning System.

A. The Design

The proposed system works in two phases; offline and online phase. The offline phase involves creating a radio map. The radio map will store distributions of RSS values from all detected APs at specific points which are known as marking positions. The marking positions together with the MAC address of each detected APs and their corresponding RSS values will be stored in the database to create the radio map.

Creating Radio Map

The radio map will consist of a set of samples taken at specific locations on a map called fingerprints. The fingerprint will consist of the location name and a measurement vector which consists of all the detected Access points and their corresponding signal strength.

The measurement vector for the signal strength will be illustrated as below;

$$a = \{a_1, a_2, \dots, a_i, a_P\}$$

Where a is the location name, a_P is the number of access points detected at this location and a_i is the signal strength of access point no. i . For instance, if a location is to be taken at the following point on this floor plan,



Fig 2: Room floor plan

The signal vector should be $104_1 = \{-30, -25, -33, -68, \dots\}$. In the database, the sample 104_1 will be connected to its corresponding location stated as:

$$104_1 = \{x_coord, y_coord, floor\}$$

Basically, the radio map will consist of a series of sample (fingerprint) which will be manually measured at specific points and stored in the database of the server. For each sample, the measurement will be taken at four different directions (north, east, south, and west). As it was said before, signal strength is very sensitive to attenuation caused by different factors such as human body, interference and furniture. Therefore, at each specific location, measurement will be taken at four different directions and the mean values of the data obtained at these directions will be used as the final measurement for the location.

Filtering Technique

A filtering technique is applied when collecting signal strength in order to reduce the number of measured signal strength that will be used to represent the fingerprint of a current location. Moreover, by using a reduced number of signal strength, the time for computation can be reduced as well as the size of storage required to store the data. At each specific location, the range of RSSI values which will be used is:

$$-90 < ss < -20$$

Interpolating Algorithm

For the interpolation technique, the same technique that was adopted by Tsai et al. will be used. By using interpolation, the

time to build the radio map will be reduced considerably. To calculate data for un-calibrated grid points, they used either one of the following formulas based on the situation. For instance, point A and B was calibrated and we need to calculate for point C which is in between, if only point A is used to infer the location of point C, the first equation is used and if both points are used to infer the location of point C, then the second equation is used. After the grid-points have been calculated, Segment process is used to divide the data of each point into m parts.

$$S_{Ci} = \frac{\log d_0}{\log (d_0 + d_1)} \times S_{Ai}$$

$$S_{Ci} = \frac{\log d_2}{\log (d_1 + d_2)} \times S_{Ai} + \frac{\log d_1}{\log (d_1 + d_2)} \times S_{Bi}$$

Eq 2: Interpolation Model [41]

Matching Algorithm

The online phase is the process where the mobile phone gets the current signal strength from detected access points and these signals are sent to the database to be matched with the stored fingerprint.

For the matching algorithm, Euclidean distance will be used to compare the current fingerprint obtained at the mobile application and the existing fingerprint stored in the database of the server. The Euclidean distance will compute the minimal distance between two set of fingerprint. For instance, assume that the current fingerprint is $s = \{s_1, s_2, \dots, s_N\}$ and a saved fingerprint is $S = \{S_1, S_2, \dots, S_N\}$, then the squared Euclidean distance between the vectors s and S is:

$$L(s, S) = (s_1 - S_1)^2 + \dots + (s_N - S_N)^2$$

This can be represented as

$$L = \sqrt{\sum_{i=1}^n |s_i - S_i|^2}$$

Eq 3: Euclidean distance algorithm [14]

User's Collaboration

Another good point of the proposed system is user's collaboration. What is meant by this is that, since it has been seen that the most inconvenient fact of Fingerprint-based Algorithm is training the system, by using the collaboration of users, the time spent on this phase can be reduced as well as increasing the accuracy of the system.

The idea behind the collaboration of the users is to let them create and manage the locations. It is obvious that to train a building of 100 rooms, it requires lots of times as well as personals. With the collaboration of the user, not all the places need to be fully trained. By using the proposed system, every user can generate, manage and, above all, use location information that was created by other users. Therefore, if a location is unknown while a user is using the system, he or she can easily update the location with an approximate name or

location so that in the future, someone else can use this location.

This feature can also help in increasing the accuracy of the location as since the user has full access on managing a location, he or she can update a location if this one is not accurate.

B. Architecture

The architecture of the proposed system is divided into three different components: The client application, the Symbian sniffer and the server. These three components are discussed further in this segment.

The Symbian Sniffer

The Symbian Sniffer is the component which will be in responsible of detecting WiFi access points and collecting the necessary information such as the MAC address, the Network name and the Signal strength of each of the access points. The component which will be coded in Symbian will be installed on the mobile phone.

It was necessary to separate the mobile application into two components because since the client application is coded in J2ME, it has some limitation of gathering network information. Therefore, the Symbian API will allow us to collect this information and send them to the J2ME application.

The Mobile Application

The mobile application which is on the client side will be developed using J2ME. This component will serve as an intermediate between the Symbian Sniffer and the Server. As it was mentioned above, the J2ME application is not able to collect network information and as a matter of fact, the Sniffer's job is to collect the necessary network information and send them to the client application.

The J2ME application will have a two way communication with the Server-side. After collecting the necessary information from the Sniffer, this information which comprises MAC addresses, network names and Signal Strength of detected access points, will be sent as a fingerprint to the server-side. This fingerprint will be compared with stored fingerprints and if a match is found, the corresponding location will be returned to the client application.

The Server

Since the mobile application is coded in J2ME, to facilitate the reusing of code, the Server is developed using J2SE and MySQL for the database. The Server provides different services such as storing fingerprint in other way to save a location in the database.

Another service is the retrieving of maps and location. The floor plan of associated location is retrieved from the web server where all the images of the floor plans are stored.

Finally, it provides service to locate a mobile phone. When a current fingerprint is sent to the server, it is compared to stored fingerprints inside the database and by using the matching algorithm, the location that best matches the measurement taken by the mobile application is retrieved from the database. Based on the selected fingerprint, the associated x and y coordinates together with the appropriate map is

returned to the j2me application. Moreover, every mobile device uses the same database of fingerprints. This allows to easily sharing knowledge about locations and enables a quick mapping of a building

IV. IMPLEMENTATION

This section describes the implementation of the different components of the proposed system. Here, the technical aspect will be explained together with some piece of codes and screenshots.

A. The coding

Symbian C++

The Symbian Sniffer was coded in Symbian C++ by using Carbide C++ as the IDE. The following portion of code is used to retrieve network name and signal strength from detected access points. The Sniffer was based on the RedpinSniffer developed by Redpin.org. [1]

```
void CWlanInfo::NetworkNameL()
{
    TPkgBuf<TConnMonNetworkNames> networks;
    RConnectionMonitor monitor;
    monitor.ConnectL();
    CleanupClosePushL(monitor);

    TRequestStatus status;

    monitor.GetPkgAttribute(EBearerIdWLAN, 0, KNetworkNames,
        networks, status); // EBearerIdWLAN KNetworkNames
    User::WaitForRequest(status);
    User::LeaveIfError(status.Int());
    // TBuf<20> ibuff; now in header max 50
    TInt count = networks().iCount;
    ibuff.Zero();
    ibuff.AppendNum(count);

    // the buffer for the file writing
    HBufC8* filebuffer = HBufC8::NewL( 200 );
    TPtr8 filebufferPtr = filebuffer->Des();
    _LIT8(KFormat1,"%S,%d");
    //open the connection to file
    RFs fs;
    fs.Connect();
    RFile file;
    if(file.Open(fs, _L("C:\\Data\\output_data.txt"))
    EFileWrite|EFileShareAny) != KErrNone)
    file.Replace(fs, _L("C:\\Data\\output_data.txt"),
    EFileWrite|EFileShareAny);
    TInt pos = 0;
    file.Seek(ESeekEnd,pos);
    for(TInt i=0;i<count;i++)
    {
        TBuf8<32> ssid;
        ssid.Copy( networks().iNetwork[i].iName );
        ibuff.Zero();
        ibuff.Copy(ssid);
        ((CSymbianSnifferAppUi*)(CEikonEnv::Static()->AppUi()))-
        >DebugEngine()->PrintLn(_L("Network Name: "));
        ((CSymbianSnifferAppUi*)(CEikonEnv::Static()->AppUi()))-
        >DebugEngine()->PrintLn(ibuff);
        TUint8 iSignal = networks().iNetwork[i].iSignalStrength;
        ibuff.Zero();
        ibuff.AppendNum(iSignal);
        ((CSymbianSnifferAppUi*)(CEikonEnv::Static()->AppUi()))-
        >DebugEngine()->PrintLn(_L("signal strength"));
```

```
((CSymbianSnifferAppUi*)(CEikonEnv::Static()->AppUi()))-
>DebugEngine()->PrintLn(ibuff);
```

```
filebufferPtr.Format(KFormat1,&ssid, iSignal);
_LIT8(KFormat2,"%n");
filebufferPtr.Append(KFormat2);
file.Write( filebufferPtr );
}
```

The next portion is for retrieving MAC Address in Symbian C++. The data structure being used is stack, where the info are stack on each other and push down each time a new data is inserted.

```
void CWlanInfo::GetMacAddressL()
{
    TBuf<10> infoBuff;
    CWlanScanInfo* scanInfo=CWlanScanInfo::NewL();
    CleanupStack::PushL(scanInfo);
    CWlanMgmtClient* client=CWlanMgmtClient::NewL();
    CleanupStack::PushL(client);
    client->GetScanResults(*scanInfo);
```

J2ME

The mobile application was built in J2ME. The function to retrieve network information had to be coded separately in Symbian because J2ME has limitation on capturing this kind of information. After the SymbianSniffer retrieves the network information, it is transferred to the mobile application. The following codes are used to establish the connection between the client application and the Symbian Sniffer.

```
private void setupConnection() {
    try {
        this.connection=(StreamConnection) Connector.open("socket://"
            + StaticResources.SNIFFER_HOST + ":"
            +String.valueOf(StaticResources.SNIFFER_PORT));
        LogService.info(this, "socket://" +
            StaticResources.SNIFFER_HOST + ":"
            +String.valueOf(StaticResources.SNIFFER_PORT));

        LogService.info(this,"setupconnection succesfully");

        /* setup Output Stream */
        this.output=new PrintStream(connection.openOutputStream());
        LogService.info(this, "setup output stream succesfully");

        /* setup Input Stream */
        this.inputStream = connection.openInputStream();
        LogService.info(this, "setup input stream succesfully");
```

Below is the function for matching the current fingerprints with the stored fingerprints. As default the number of matches is 0. This number 0 will increase based on the number of similar fingerprints which have been found during the comparison stage.

```
Vector WiFiReadings1 = this.getWiFiReadings().getVector();
Vector WiFiReadings2 = m.getWiFiReadings().getVector();
matches = 0;
for (int i = 0; i < WiFiReadings1.size(); i++) {
    WiFiReading WiF1 = (WiFiReading) WiFiReadings1.elementAt(i);
    for (int j = 0; j < WiFiReadings2.size(); j++) {
        WiFiReading WiF2 = (WiFiReading)
        WiFiReadings2.elementAt(j);
```


B. The mobile application in action

This subdivision illustrates the complete process of the application in finding the location of the user together with some screenshots of the application in action. This part shows exactly what will happen on the user's side.

Initialization

The initialization process or loading process is the most important process of the application as most of the functions happen here. The first function is that the application is going to check and retrieve data stored in the preferences such as the server name and addresses.

The next step is to set and verify if the connection with the Sniffer is established. If the connection is established, the application will move to the next phase which will be to retrieve the network information from the Symbian Sniffer. This phase is known as scanning radios.

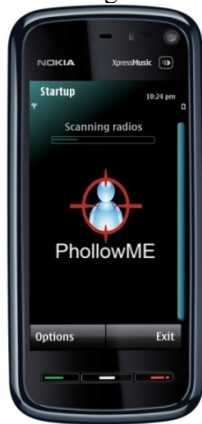


Fig 3: Picture showing the phase of scanning radios

After the information has been passed to the J2ME application, this information which is represented as a fingerprint will be sent to the server to be compared to stored fingerprint inside the database. This phase will be identified on the screen by "Retrieving position".



Fig 4: Location is indicated by a red crosshair.

If a match is found, the server will return the selected fingerprint together with its associated x and y coordinates and

the appropriate map (floor plan). If in case, no match is found, the application will return a message stating that the location is unknown and should be updated. This update process allows the user to name the unknown location and show approximately its location on a floor plan. The following picture a known location which is indicated on a floor plan by a red crosshair.

Updating stage

If a location is unknown, the application will suggest the user to update the location. The stage consists of 3 sections which is one to give a name to the location e.g Room 44; two is to select the appropriate floor plan and three is to select the location by moving the crosshair over it.

The first step of the update stage is to name the location. A textfield will be presented to the user to enter any name using any character to represent the location. This name will be saved in the database together with the collected measurement as a fingerprint.

After a name has been entered to represent the location, the user will be presented a list of available maps from which he or she will have to choose the appropriate one. For instance, if the user knows that the location where he or she is standing is situated on the third floor of the building, the user needs to choose the floor plan of the third floor.

Then, the floor plan will appear on the screen of the mobile phone together with a red crosshair which the user will be able to move. The user will have to move the crosshair approximately where he or she is standing in order for the system to gather the x and y coordinates of the location. The user will be able to know where he is if he or she knows the location name as each room on the floor plan will be labeled.

Finally, by saving the location, the current fingerprint together with the selected map and the x and y coordinates of the location will be stored in the database.

V. TESTING

A. Testing the system

The testing of the proposed system was performed in a university. Only on floor plan was used during the testing. On the floor plan, there were 12 rooms and 6 access points placed at different location. For the testing only three of the total rooms were trained. Two of the trained rooms were located next to each other whereas the third one was alone quite far from the two.

The training of the location was performed using the proposed system itself. A laptop was also used to compare the results from time to time. The software which was used on the laptop was "WirelessMon".

By using the proposed system to train the locations also demonstrates the efficiency of the Update function of the system. If in case, a location was unknown during the initialization stage, the following message was presented to the user to state that the location is unknown in the database and has to be updated. When updating a location, it is preferable to locate the cross hair in the middle of the room. This actually avoid measurements that are similar but that are

taken from different rooms, for instance if the rooms are next to each other.

When the training was done, there was only the user inside the room. To check the accuracy of the application, when the application was used during on-line phase, we adopted the same environmental situation, that is an empty room and the results obtained were satisfactory. We also try to find our location in a different circumstance where the same room was full of students and the result that was obtained was quite inaccurate.

In one case, it was shown that the location was unknown as for this room, only one measurement was taken. To solve this problem, more measurements have to be taken during the off-line phase.

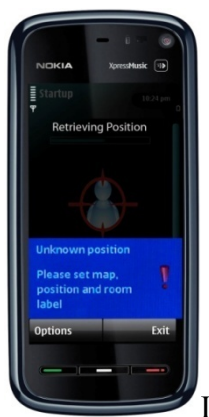


Fig 5: Location is unknown

B. Performance Evaluation

The section expresses the performance of the proposed system in terms of time taken in performing a function. Some of the functions that have been tested are the time taken in storing a fingerprint, in retrieving a location and in retrieving a map from the web server. The compression size of file is also illustrated in this section.

Storing a Fingerprint

The function of storing a fingerprint occurs during the update phase. Here, the current measurement together with the associated map and coordinates of the location will be stored in the database.

This process takes around 5-6 seconds because it has to read the measurements, serialized it and sends it to the server. At the server side, the serialized object will be de-serialized to return to its original format to be stored in the database. The serialization of data was necessary to accelerate the process as the all the data is compressed before being sent to the server.

Retrieving a location

The process of retrieving a location is where the current fingerprint is compared to the stored fingerprints and if a match is found, the location is returned to the mobile phone. From the test which has been performed, the result is relatively good in terms of time. The amount of time taken by this process is approximately .075 seconds.

The amount of time for retrieving location need to be short so that when the user initialized the system, the location is returned in a reasonable time.. Moreover, if the user is moving while searching his position, the time to get location should be fast so that the position of the user on the screen follows the real position of the user since the refresh rate is once every 2 seconds.

Retrieving a Map

Based on matching fingerprint, the relative map is fetched from the web server to be displayed on the mobile screen. In the tests performed, the retrieval of map is done in 0.050 second to 0.070 second.

Serialization of Data

For this test, a dummy measurement was compressed to check the process of serialization. The average size of compression is between 325 – 400 kb. This shows that the file that will be transferred to the server is reasonable small thus reducing the amount of time for the localization process. The amount of time for compressing the data is approximately of 0.955 second to 1.10 second.

Accuracy

The accuracy of the application was hard to define precisely. It has been shown that many factors can affect the accuracy such as for instance, the number of fingerprint sample taken for each room. Actually, the more fingerprints taken for a room, the more precise will the location be. Another point that affected the accuracy is the amount of people in the room. Two tests were performed where measurement were taken in a room at two different time, one when the class was empty and one when the class was full. The result obtained from the tests was not similar and the difference between the locations was 1.71m. The reason was that when the system was trained, the room was empty but when the test was performed in a room full of student, the result was spoiled.

Since the application depends on users' collaboration, it is difficult to know under which circumstances, the measurements are taken during the updating phase; it is preferable when updating a location to be standing at the center of the room. By doing so, we assure that the measurement taken is not close to the measurement that is taken in the neighbor room. In a way, we increase the accuracy of the location.

In order to have an accurate system, the main objective is to train the system to its maximum. The more measurement we have, the more precise will be the matching process. Unfortunately, the time that would be spent to train a system fully is not reasonable mostly if the building is huge.

VI. ADVANTAGES OF THE SYSTEM

First of all, the proposed system will be using WIFI signal strength to estimate the position of the user. WIFI was chosen as the signal source because it has been found to be the most economical technology to be used for the implementation of the proposed system.

The approach used for developing the system was a Fingerprint-based technique. Fingerprint-based has been proved to be more accurate than using other technique such as

mathematical propagational model or triangulation. By using Fingerprint approach, the proposed system is built into two phases, offline phase where the system is trained and online phase where the user uses the mobile application to infer his position.

Most of the improvements were done in the offline phase. Basically, many researchers avoid using fingerprint approach because of the offline phase. Taking measurements at various locations in a building is very laborious and usually discourage developers to adopt this technique. The strength of the proposed system has revised this phase and uses techniques to improve the accuracy and also reduce the amount of time spent on calibrating the system. One of the improvements was done at the process of collecting fingerprints. To avoid inaccuracy of signal strength caused by attenuation due to human body or other factors, at each marking position, signal strength was taken at four different directions. Also, a filtering technique was used to take signal strength that lies between the range of -90 and -20 dBm. This filtering was used to discharge needless signal strength and thus reducing the time for computation and reducing the size of storage for the fingerprints. Moreover, to reduce the time of calibration, I have used an interpolation technique. By using this technique the amount of fingerprint that needs to be manually measured has been divided by half.

Finally, to infer the position of the user, the matching algorithm that was chosen is the Euclidean distance. Euclidean distance was used to compute the minimal distance between the current fingerprint and the existing fingerprint.

VII.CONCLUSION

Nowadays, positioning system is very useful in outdoor environment as well as indoor environment. Indoor environment is increasing in size and is becoming more complex. Therefore, developing an indoor positioning system is indispensable as it will avoid stress and reduce time for people to look for a specific location in an indoor environment.

Also, since we are moving to ubiquitous computing and that technologies are increasing, what researchers are visioning is to make mobile phone not only a communication tool but also a navigation tool. To conclude, the system that was proposed and developed in this major project is still in the stage of prototype. The intention of this system is to collaborate in the research on indoor positioning system and to do one more step in building an indoor positioning system that will be feasible and used in real-life.

VIII.FUTURE WORKS

As a future enhancement, the accuracy of the system can be increased by applying various techniques such as backtracking which will avoid the system from choosing between two close measurements. By applying backtracking, when ambiguity occurs, the system will go back to the user's previous location

and then compare which might be the user's next location between the candidate locations.

Another future improvement will be to use other signal source such as Bluetooth and GSM wireless technologies. By having more signal sources, the accuracy of the system will increase and the system will be able to be used on bigger scale.

The main goal is to make the system feasible to be used in real-life. Until now, it is still in the prototype phase where various improvements need to be amended and research need to be perform on how to integrate the system in real-life. The main objective is on how to get a signal that is less sensitive to attenuation and to develop a system that can be implemented in different indoor environment.

ACKNOWLEDGMENT

A sincere gratitude and appreciation is dedicated to Mr. Tee Wee Jing for his contribution in this project and a special consideration goes to our family for their continuous support and encouragement. Also, the special thank goes to our helpful advisor Dr. Arash Habibi Lashkari for his advising and guidance in the progression of our dissertation and publication.

REFERENCES

- [1] History of GPS. Retrieved from <http://en.wikipedia.org/wiki/GPS#History> on 26th July 2009.
- [2] M. Garcia, C. Martinez, J. Tomas and J. Lloret, "Wireless Sensors self-location in an Indoor WLAN environment", International Conference on Sensor Technologies and Applications, pp. 146-151 (2007)
- [3] R. Want, A. Hopper, V. Falcao and J. Gibbons, "The Active Badge Location System", Olivetti Research Ltd, England. Retrieved from <http://www.cl.cam.ac.uk/research/dtg/publications/public/files/tr.92.1.pdf>
- [4] J.Korhonen, T. Ojala, M. Klemola, and P. Vaanallen, "mTag- Architecture for discovering Location Specific Mobile Web Services Using RFID and Its Evaluation with Two Case Studies".
- [5] P. Bahl and V. N. Padmanabhan, "RADAR: An RF-Based In-Building User Location and Tracking System", IEEE INFOCOM, March 2000
- [6] Herecast: WiFi Location based services/802.11 Positioning System. Retrieved from <http://www.herecast.com>
- [7] Shaun Phillips, Michael Katchabaw, Hanan Lutfiyya, "WLocator: An Indoor Positioning System," wimob, pp.33, Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2007), 2007
- [8] W. H. Wong, J. K. Ng and W. M. Yeung. Wireless LAN Positioning with Mobile Devices in a Library Environment. In Proceedings of the 25th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW'05) 2005
- [9] K. Kaemarungsi and P. Krishnamurthy. Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting. In proceedings of the first Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous' 04), 2004
- [10] P. Bahl and V. N. Padmanabhan, A Software System for Locating Mobile Users: Design, Evaluation and Lessons. University of California, San Diego
- [11] D.-B Lin R.-T Juang and H.-P Lin. Robust Mobile Location Estimation Based on Signal Attenuation for Cellular Communication Systems. National

Taipei University of Technology, Institute of Computer and Communication,
Taipei, Taiwan, Republic of China (2004)

[12] Caffery Jr., J.: A New Approach to the Geometry of TOA Location. In: Proc. IEEE Vehicular Technology Conference (VTC 2000-Fall), September 2000, vol. 4, pp. 1943-1949 (2000).

[13] S. A. Golden and S. S. Bateman. Sensor Measurements for WIFI Location with Emphasis on Time-of-Arrival Ranging, IEEE Transactions on Mobile Computing, Vol 6, NO 10, October 2007

[14] A. Nafarieh and J. Ilow, A Testbed for Localizing Wireless LAN Devices Using Received Signal Strength, Communication Networks and Services Research Conference, 2008

Kekre's Wavelet Transform for Image Fusion and Comparison with Other Pixel Based Image Fusion Techniques

Dr. H.B. kekre
MPSTME, SVKM'S
NMIMS university
hbkekke@yahoo.com

Dr.Tanuja Sarode
Computer engineering department,
Thadomal Shahani Engineering college
tanuja_0123@yahoo.com

Rachana Dhannawat
Computer Sci. & engg. department,
S.N.D.T. University, Mumbai.
rachanadhannawat82@gmail.com

ABSTRACT- Image fusion combines several images of same object or scene so that the final output image contains more information. The main requirement of the fusion process is to identify the most significant features in the input images and to transfer them without loss into the fused image. In this paper many pixel level fusion techniques like DCT averaging, PCA, Haar wavelet and Kekre's wavelet transform techniques for image fusion are proposed and compared. The main advantage of Kekre's transform matrix is that it can be of any size $N \times N$, which need not to be an integer power of 2. From $N \times N$ Kekre's transform matrix, we can generate Kekre's Wavelet transform matrices of size $(2N) \times (2N)$, $(3N) \times (3N)$,, $(N^2) \times (N^2)$.

I. INTRODUCTION:

Image fusion is the technology that combines several images of the same area or the same object under different imaging conditions. In other words, it is used to generate a result which describes the scene "better" than any single image with respect to relevant properties; it means the acquisition of perceptually important information. The main requirement of the fusion process is to identify the most significant features in the input images and to transfer them without loss of detail into the fused image. The final output image can provide more information than any of the single images as well as reducing the signal-to-noise ratio.

The object of image fusion is to obtain a better visual understanding of certain phenomena, and to enhance intelligence and system control functions. Applications of image fusion might use several sensors like thermal sensor, sonar, infrared, Synthetic Aperture radar (SAR), electro-optic imaging sensors Ground Penetrating Radar (GPR), Ultra Sound Sensor (US), and X-ray sensor. The data gathered from multiple sources of acquisition are delivered to preprocessing such as denoising and image registration. This step is used to associate the corresponding pixels to the same physical points on

the object. In this method, the input images can be compared pixel by pixel. The post-processing is applied to the fused image. Post-processing includes classification, segmentation, and image enhancement.

Many image fusion techniques pixel level, feature level and decision level are developed. Examples are like Averaging technique, PCA, pyramid transform [7], wavelet transform, neural network, K-means clustering, etc.

Several situations in image processing require high spatial and high spectral resolution in a single image. For example, the traffic monitoring system, satellite image system, and long range sensor fusion system, land surveying and mapping, geologic surveying, agriculture evaluation, medical and weather forecasting all use image fusion.

Like these, applications motivating the image fusion are:

1. Image Classification
2. Aerial and Satellite imaging
3. Medical imaging
4. Robot vision
5. Concealed weapon detection
6. Multi-focus image fusion
7. Digital camera application
8. Battle field monitoring

II. PIXEL LEVEL FUSION TECHNIQUES:

1) *Averaging Technique* [4]:

This technique is a basic and straight forward technique and fusion could be achieved by simple averaging corresponding pixels in each input image as

$$F(m,n) = (A(m,n) + B(m,n)) / 2 \quad (1)$$

The simplest way to fuse two images is to take the mean-value of the corresponding pixels. For some applications this may be enough, but there will always be one image with poor lighting and thus the

quality of an averaged image will obviously decrease. Averaging doesn't actually provide very good results.

2) Principal Components Analysis [8]:

Principal component analysis PCA is a general statistical technique that transforms multivariate data with correlated variables into one with uncorrelated variables. These new variables are obtained as linear combination of the original variables. It is used to reduce multidimensional data sets to lower dimensions for analysis. The implementation process may be summarized as:

- (i) Take as input two images of same size.
- (ii) The input images (images to be fused) are arranged in two column vectors;
- (iii) The resulting vector has a dimension of $n \times 2$, where n is length of the each image vector; Compute the eigenvector and eigen values for this resulting vector and the eigenvectors corresponding to the larger eigen value obtained, and
- (iv) Normalize the column vector corresponding to the larger Eigen value.
- (v) The values of the normalized Eigen vector act as the weight values which are respectively multiplied with each pixel of the input images.
- (vi) Sum of the two scaled matrices calculated in (vi) will be the fused image matrix.

The fused image is:

$$I_f(x,y) = P_1 I_1(x,y) + P_2 I_2(x,y) \quad (2)$$

Where P_1 and P_2 are the normalized components and its equal to $P_1 = V(1) / \sum V$ and $P_2 = V(2) / \sum V$ where V is eigen vector and $P_1 + P_2 = 1$.

3) Discrete Cosine Transform Technique:

Discrete cosine transform (DCT) is an important transform in image processing. An image fusion technique is presented based on average measure defined in the DCT domain. Here we transform images using DCT technique and then apply averaging technique finally take the inverse discrete cosine transform to reconstruct the fused image. Actually, this image fusion technique is called the DCT + average; modified or "improved" DCT technique [5] as shown in figure 2.1.

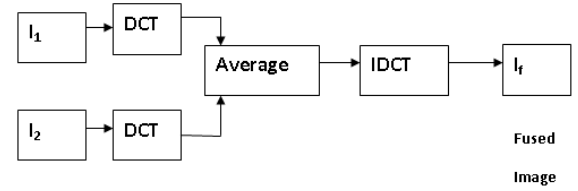


Fig. 2.1. Schematic diagram for the DCT based pixel level image fusion scheme

4) Discrete Wavelet Transform Technique with Haar based fusion:

With wavelet multi-resolution analysis [2] and fast Mallet's transform [1], the algorithm first decomposes an image to get an approximate image and a detail image, which respectively represent different structures of the original image i.e. the source images A and B are decomposed into discrete wavelet decomposition coefficients: LL (approximations), LH, HL and HH (details) at each level before fusion rules are applied. The decision map is formulated based on the fusion rules. The resulting fused transform is reconstructed to fused image by inverse wavelet transformation and Wavelet transform has the ability of reconstructing, so there is no information loss and redundancy in the process of decomposition and reconstruction. The fast Mallet's transform largely decreased the time of operation and made its application possible in image processing.

The wavelet transform is based on the orthogonal decomposition of the image onto a wavelet basis in order to avoid a redundancy of information in the pyramid at each level of resolution, the high and low frequency components of the input image can be separated via high-pass and low-pass filters. Thus, the image fusion with the wavelet multi-resolution analysis can avoid information distortion; ensure better quality and showing more spatial detail. Therefore, comparing with other methods such as averaging, DCT, pyramid and PCA, the wavelet transform method has better performance in image fusion.

The Haar wavelet is the first known wavelet.

The 2×2 Haar matrix that is associated with the Haar wavelet is

$$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (3)$$

4x4 Haar transformation matrix is shown below.

$$H_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} \dots\dots(4)$$

4) Kekre's Transform:

Kekre's transform matrix [11] can be of any size NxN, which need not to be an integer power of 2. All upper diagonal and diagonal elements of Kekre's transform matrix are 1, while the lower diagonal part except the elements just below diagonal is zero. Generalized NxN Kekre's transform matrix can be given as,

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ -N+1 & 1 & 1 & \dots & 1 & 1 \\ 0 & -N+2 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & -N+(N-1) & 1 \end{bmatrix} \quad (5)$$

The formula for generating the element Kxy of Kekre's transform matrix is,

$$K_{xy} = \begin{cases} 1 & : x \leq y \\ -N + (x - 1) & : x = y + 1 \\ 0 & : x > y + 1 \end{cases} \quad (6)$$

Kekre's Wavelet Transform [6]:

Kekre's Wavelet transform is derived from Kekre's transform. From NxN Kekre's transform matrix, we can generate Kekre's Wavelet transform matrices of size (2N)x(2N), (3N)x(3N),....., (N²)x(N²). For example, from 5x5 Kekre's transform matrix, we can generate Kekre's Wavelet transform matrices of size 10x10, 15x15, 20x20 and 25x25. In general MxM Kekre's Wavelet transform matrix can be generated from NxN Kekre's transform matrix, such that M = N * P where P is any integer between 2 and N that is, 2 ≤ P ≤ N. Consider the Kekre's transform matrix of size NxN shown in fig. 2.2.

K ₁₁	K ₁₂	K ₁₃	...	K _{1(N-1)}	K _{1N}
K ₂₁	K ₂₂	K ₂₃	...	K _{2(N-1)}	K _{2N}
K ₃₁	K ₃₂	K ₃₃	...	K _{3(N-1)}	K _{3N}
⋮	⋮	⋮	⋮	⋮	⋮
K _{N1}	K _{N2}	K _{N3}	...	K _{N(N-1)}	K _{NN}

Fig. 2.2 Kekre's Transform (KT) matrix of size NxN

Fig. 2.4 shows MxM Kekre's Wavelet transform matrix generated from NxN Kekre's transform matrix. First N numbers of rows of Kekre's Wavelet transform matrix are generated by repeating every column of Kekre's transform matrix P times. To generate remaining (M-N) rows, extract last (P-1) rows and last P columns from Kekre's transform matrix and store extracted elements in to temporary matrix say T of size (P-1) x P. Fig.2.3 shows extracted elements of Kekre's transform matrix stored in T.

K _{(N-P+2)(N-P+1)}	K _{(N-P+2)(N-P+2)}	...	K _{(N-P+2)N}
K _{(N-P+3)(N-P+1)}	K _{(N-P+3)(N-P+2)}	...	K _{(N-P+3)N}
⋮	⋮	⋮	⋮
K _{N(N-P+1)}	K _{N(N-P+2)}	...	K _{NN}

Fig. 2.3 Temporary matrix T of size (P-1) x P

1st column of kekre's transform
repeated P times

2nd column of kekre's transform
repeated P times

Nth column of kekre's transform
repeated P times

K_{11}	K_{11}	...	K_{11}	K_{12}	K_{12}	...	K_{12}	...	K_{1N}	K_{1N}	...	K_{1N}
K_{21}	K_{21}	...	K_{21}	K_{22}	K_{22}	...	K_{22}	...	K_{2N}	K_{2N}	...	K_{2N}
K_{31}	K_{31}	...	K_{31}	K_{32}	K_{32}	...	K_{32}	...	K_{3N}	K_{3N}	...	K_{3N}
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$
K_{N1}	K_{N1}	...	K_{N1}	K_{N2}	K_{N2}	...	K_{N2}	...	K_{NN}	K_{NN}	...	K_{NN}
$K_{(N-P+2)(N-P+1)}$	$K_{(N-P+2)(N-P+2)}$	...	$K_{(N-P+2)N}$	0	0	...	0	...	0	0	...	0
0	0	...	0	$K_{(N-P+2)(N-P+1)}$	$K_{(N-P+2)(N-P+2)}$	...	$K_{(N-P+2)N}$	...	0	0	...	0
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$
0	0	...	0	0	0	...	0	...	$K_{(N-P+2)(N-P+1)}$	$K_{(N-P+2)(N-P+2)}$	...	$K_{(N-P+2)N}$
$K_{(N-P+3)(N-P+1)}$	$K_{(N-P+3)(N-P+2)}$	...	$K_{(N-P+3)N}$	0	0	...	0	...	0	0	...	0
0	0	...	0	$K_{(N-P+3)(N-P+1)}$	$K_{(N-P+3)(N-P+2)}$	...	$K_{(N-P+3)N}$	...	0	0	...	0
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$
0	0	...	0	0	0	...	0	...	$K_{(N-P+3)(N-P+1)}$	$K_{(N-P+3)(N-P+2)}$	...	$K_{(N-P+3)N}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$
$K_{N(N-P+1)}$	$K_{N(N-P+2)}$	...	K_{NN}	0	0	...	0	...	0	0	...	0
0	0	...	0	$K_{N(N-P+1)}$	$K_{N(N-P+2)}$	...	K_{NN}	...	0	0	...	0
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$
0	0	...	0	0	0	...	0	...	$K_{N(N-P+1)}$	$K_{N(N-P+2)}$	...	K_{NN}

Figure 2.4 Kekre's Wavelet transform (KWT) matrix of size MxM generated from Kekre's transform (KT) matrix of size NxN. Where $M = N * P$, $2 \leq P \leq N$.

III. PROPOSED METHOD:

1. Take as input two images of same size and of same object or scene taken from two different sensors like visible and infra red images or two images having different focus.
2. If images are colored separate their RGB planes to perform 2D transforms.
3. Perform decomposition of images using different transforms like DCT, wavelet and Kekre's Wavelet transform, etc.
4. Fuse two image components by taking average.
5. Resulting fused transform components are converted to image using inverse transform.
6. For colored images combine their separated RGB planes.
7. Compare results of different methods of image fusion using various measures like entropy, standard deviation, mean, mutual information, etc.

IV. PERFORMANCE EVALUATION IN IMAGE FUSION [3]:

At present, the image fusion evaluation methods can mainly be divided into two categories, namely, subjective evaluation methods and objective evaluation methods.

Subjective evaluation method is, directly from the testing of the image quality evaluation, a simple and intuitive, but in man-made evaluation of the quality there will be a lot of subjective factors affecting evaluation results. An objective evaluation methods commonly used are: mean, variance, standard deviation, average gradient, information entropy, mutual information and so on.

1) Standard deviation:

The standard deviation of gray image reflects clarity and contrast, the greater the value is, the higher clarity and contrast the image have; on the other hand, the smaller the image contrast is, the more affected by noise. The standard deviation is given by:

$$\sigma_x = \sqrt{\frac{1}{MN-1} \sum_{i=1}^M \sum_{j=1}^N (x(i,j) - \bar{x})^2} \quad (7)$$

Where $M \times N$ is the size of image x , $x(i, j)$ is the gray value of pixel (i, j) , $\bar{x}$ denote the mean of x .

2) Information entropy:

Information entropy [12] is an important measure of image information richness, which indicates the average information amount contained in the image. The greater of the entropy is the greater of the amount of information carried by the fusion image. Based on gray-scale L and the gray distribution probability p_i of pixels, then the image entropy is given as follows:

$$H = -\sum p_i \log(p_i) \quad (8)$$

3) Mean:

Mean gray image reflects the image brightness, the greater of the mean gray is, the higher of the brightness. However, the brightness of the image is not necessarily as high as possible; usually in the median low of the gray-scale range have a better visual effect.

4) Mutual Information:

The mutual information is often used for fusion evaluation. Mutual information [10] of image A and F can be defined as:

$$I(x_A, x_F) = H(x_A) + H(x_F) - H(x_A, x_F) \quad (9)$$

Where $H(x_A)$ is the entropy from image 1, $H(x_F)$ is the entropy from image 2, and $H(x_A, x_F)$ is the joint entropy. The measure $I(x_A, x_F)$ indicates how much information the composite image x_F conveys about the source image x_A . Thus, the higher the mutual information between x_F and x_A , the more x_F

resembles the image x_A . In this sense, mutual information can be interpreted as a 'similarity' measure. Consider two input image, a measure based on mutual information proposed by Gema[9], that is obtained by adding the mutual information between the composite image and each of the inputs, and dividing it by the sum of the entropies of the inputs, i.e.,

$$MI(x_A, x_B, x_F) = (I(x_A, x_F) + I(x_B, x_F)) / (H(x_A) + H(x_B)) \quad (10)$$

The higher the value in (9), the better the quality of the composite image is supposed to be.

V. RESULTS and ANALYSIS:

Above mentioned techniques are tried on pair of three color RGB images and six gray images as shown in fig 5.1 and results are compared based on measures like entropy, mean, standard deviation and mutual information [3]. Figure 5.2 shows Image fusion by different techniques for visible and infra red scenery images. Figure 5.3 shows Image fusion by different techniques for hill images with different focus. Figure 5.4 shows Image fusion by different techniques for gray clock images with different focus. Figure 5.5 shows Image fusion by different techniques for gray ct and mri medical images. Performance evaluation based on above mentioned four measures for color image is given in table 5.1. Table 5.2 presents performance evaluation for gray images.



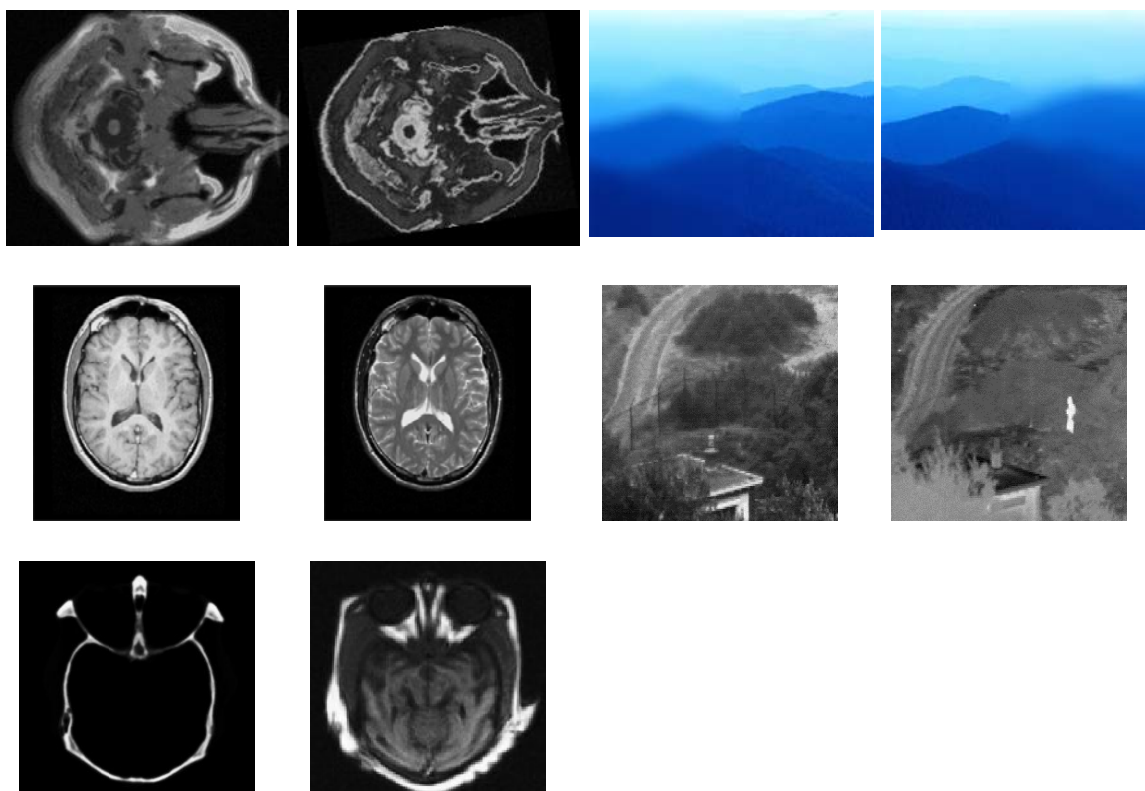


Fig. 5.1: Sample images



Fig. 5.2 Image fusion by different techniques for visible and infra red scenery images

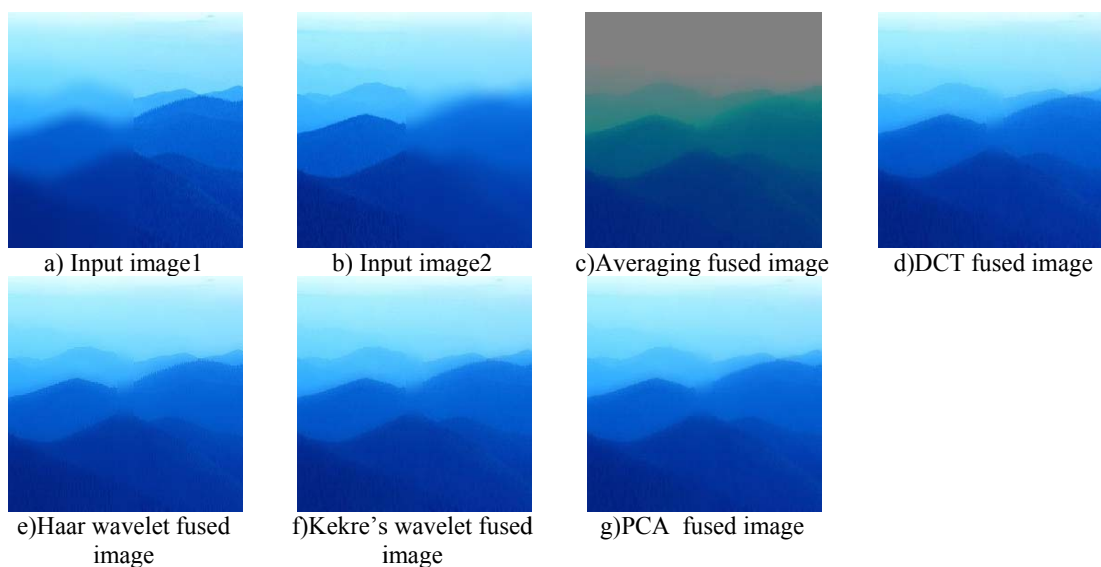


Fig. 5.3 Image fusion by different techniques for hill images with different focus

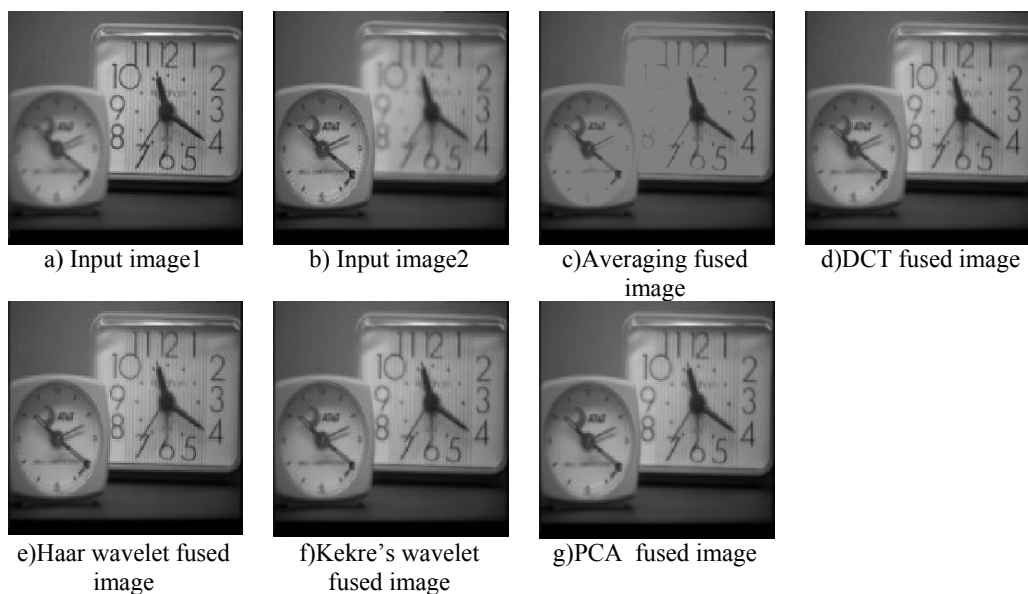
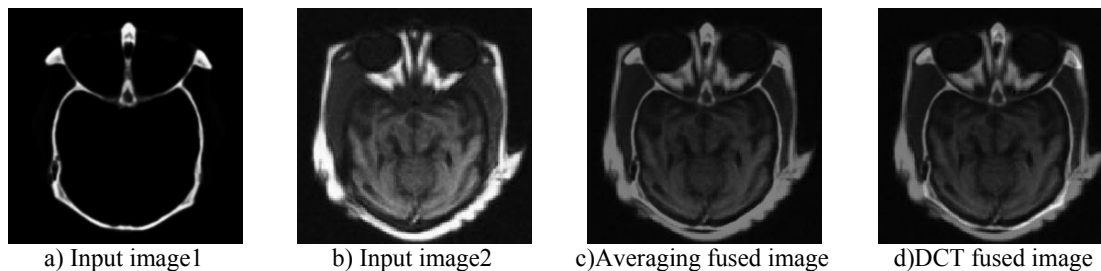


Fig. 5.4 Image fusion by different techniques for clock images with different focus



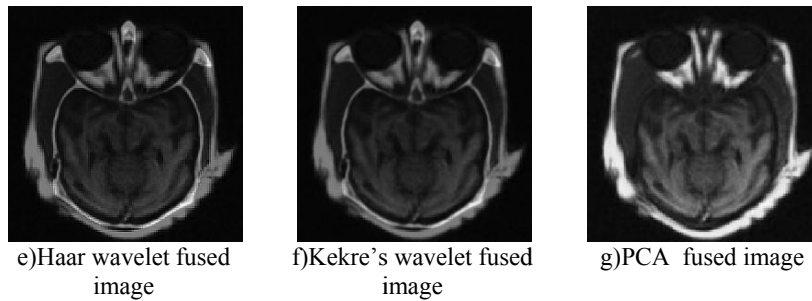


Fig. 5.5 Image fusion by different techniques for ct and mri images

Table 5.1 Performance evaluation for color images

		Averaging	DCT	PCA	Haar wavelet	Kekre's wavelet
Scenery image	Mean	74.0107	88.6090	91.6637	88.9377	88.8765
	SD	41.5931	64.3474	69.9428	64.5921	64.3860
	Entropy	5.6304	7.4882	7.4915	7.5192	7.4905
	MI	0.2573	0.3619	0.3781	0.3305	0.3651
Hill Image	Mean	90.7652	134.1505	134.3259	134.3870	134.4092
	SD	49.6320	90.2325	90.3185	90.3282	90.2632
	Entropy	3.6091	7.2593	7.2654	7.3650	7.2610
	MI	0.3465	0.4836	0.4892	0.4693	0.4849

Table 5.2 Performance evaluation for gray images

		Averaging	DCT	PCA	Haar wavelet	Kekre's wavelet
Clock image	Mean	89.5221	96.3092	96.4922	49.5519	96.4766
	SD	40.6857	48.9355	48.9555	49.3393	49.0089
	Entropy	4.9575	5.1872	5.1890	5.2598	5.2020
	MI	0.4316	0.5185	0.5202	0.4954	0.5182
CT MRI images	Mean	32.1246	32.2862	51.9930	32.5318	32.4113
	SD	32.7642	34.8291	53.4098	36.0796	34.8212
	Entropy	5.7703	5.9090	6.5409	5.9799	5.9108
	MI	0.5744	0.5674	0.7256	0.3982	0.5541

In table 5.1 it is observed that for scenery images mean, SD and MI is maximum by PCA technique meaning that brightness, clarity, contrast and quality of fused image is better. While entropy is maximum by Haar technique meaning that greater amount of information is carried by the fused image. For hill images mean, SD and entropy is maximum by Haar technique meaning that brightness, clarity, contrast and amount of information is carried by the fused image is more. While MI is maximum by PCA technique meaning that quality of fused image is better by this technique.

In table 5.2 it is observed that for clock images mean and MI is maximum by PCA technique meaning that brightness and quality of fused image is better. While SD and entropy is maximum by Haar technique meaning that clarity, contrast and amount of information carried by the fused image is greater. For CT and MRI images mean, SD, entropy and MI is maximum by PCA technique meaning that brightness, clarity, contrast, amount of information carried by the fused image and quality of fused image is best by this

technique. In all these images if we observe the output of the Kekre's wavelet technique it is very close to the output and the major advantage of the matrix is that it can be used for images which are not integral power of 2.

IV.CONCLUSION:

In this paper many pixel level techniques like averaging, PCA, DCT, Haar wavelet and Kekre's wavelet technique are implemented and their results are compared. It is observed that the new Kekre's wavelet transform when used for image fusion gives comparatively good results, just closer to the best result and the added advantage is that it can be used for images of any size, not necessarily integer power of 2.

REFERENCES:

- [1] Nianlong Han; Jinxing Hu; Wei Zhang, "Multi-spectral and SAR images fusion via Mallat and A trous wavelet transform ",18th International Conference on Geoinformatics, 09 September 2010, page(s): 1 - 4

- [2] Xing Su-xia, CHEN Tian-hua, LI Jing-xian "Image Fusion based on Regional Energy and Standard Deviation", 2nd International Conference on Signal Processing Systems (ICSPS), 2010, Page(s): 739 -743
- [3] Xing Su-xia, Guo Pei-yuan and Chen Tian-hua," Study on Optimal Wavelet Decomposition Level in Infrared and visual Light Image Fusion", International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), 2010 , page(s): 616 – 619
- [4] Le Song, Yuchi Lin, Weichang Feng, Meirong Zhao "A Novel Automatic Weighted Image Fusion Algorithm", International Workshop on Intelligent Systems and Applications, ISA ,2009 , Page(s): 1 – 4
- [5] MA. Mohamed and R.M El-Den" Implementation of Image Fusion Techniques for Multi-Focus Images Using FPGA" 28th National Radio Science Conference (NRSC 2011) April 26-28, 2011, Page(s): 1 – 11
- [6] Dr. H. B. Kekre, Archana Athawale, Dipali Sadavarti,"Algorithm to Generate Kekre's Wavelet Transform from Kekre's Transform" , International Journal of Engineering Science and Technology, Vol. 2(5), 2010, page(s): 756-767.
- [7] Shivsubramani Krishnamoorthy, K.P.Soman, "Implementation and Comparative Study of Image Fusion Algorithms", International Journal of Computer Applications, Volume 9– No.2, November 2010, page(s): 25-35.
- [8] V.P.S. Naidu and J.R. Raol," Pixel-level Image Fusion using Wavelets and Principal Component Analysis", Defence Science Journal, Vol. 58, No. 3, May 2008, Page(s): 338-352.
- [9] Gema Piella Fenoy, "Adaptive Wavelets and their Applications to Image Fusion and Compression", PhD thesis, Lehigh University, Bethlehem, Philadelphia, April 2003.
- [10] Li M ing-xi, Chen Jun, " A method of Image Segmentation based on Mutual Information and threshold iteration on multi-pectral Image Fusion", page(s): 385- 389.
- [11] Dr. H. B.Kekre, Dr. Tanuja K. Sarode, Sudeep Thepade, Sonal Shroff, "Instigation of Orthogonal Wavelet Transforms using Walsh, Cosine, Hartley, Kekre Transforms and their use in Image Compression", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 6, 2011, Page(s):125-133.
- [12] Koen Frenken , "Entropy statistics and information theory", July 2003.

AUTHORS PROFILE:

Dr. H. B. Kekre: has received B.E. (Hons.) in Telecomm. Engineering. From Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. For 13 years



he was working as a professor and head in the Department of Computer Engg. At Thadomal Shahani Engineering. College, Mumbai. Now he is Senior Professor at MPSTME, SVKM's NMIMS. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E./ B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 270 papers in National / International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE Recently 11 students working under his guidance have received best paper awards. Two of his students have been awarded Ph. D. from NMIMS University. Currently he is guiding ten Ph.D. students.

Dr. Tanuja K. Sarode: has Received Bsc.(Mathematics)from Mumbai University in 1996, Bsc.Tech.(Computer Technology) from Mumbai University in 1999, M.E. (Computer Engineering) degree from Mumbai University in 2004, Ph.D. from Mukesh Patel School of



Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 12 years of experience in teaching. Currently working as Assistant Professor in Dept. of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is life member of IETE, member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT), Singapore. Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has more than 100 papers in National /International Conferences/Journal to her credit.

Rachana Dhannawat: has received B.E. degree from Sant Gadg ebaba Amaravati University in 2003. She is pursuing M.E. from Mumbai University. She has more than 8years of experience in teaching. Currently working as assistant professor in Usha Mittal Institute of Technology, S.N.D.T. Univesity, Mumbai. She is life member of ISTE. Her area of interest are Image Processing,Networking, Computer graphics and algorithms.



A Survey on Building Intrusion Detection System Using Data Mining Framework

V. Jaiganesh, Assistant Professor
Department of Computer Science
Dr.N.G.P. Arts and Science College
Coimbatore
e-mail: jaiganeshree@gmail.com

M. Thenmozhi Assistant Professor
Department of Information
Technology
Avinashilingam University for
Women, Coimbatore
e-mail: thenujai@gmail.com

Dr. P. Sumathi, Assistant Professor
Department of Computer Science,
Chikkanna Government Arts College,
Tirupur
e-mail: sumi_rajes@yahoo.com

Abstract— Recently, network attacks have increased to a greater extent. Hackers and intruders can produce several successful efforts to cause the crash of the networks and web services by illegal intrusion. New threats and interrelated solutions to avoid these threats are budding jointly with the secured system evolution. So, Intrusion Detection System (IDS) has become an active area of research in the field of network security. The optimization of IDS becomes an attractive domain due to the security audit data as well as complex and active properties of intrusion behaviors. The main purpose of IDS is to protect the resources from threats. Intrusion Detection System examines and calculates the user behavior, and then these behaviors will be considered an attack or a normal behavior. Intrusion detection systems have been integrated with data mining approaches to identify intrusions. There are various data mining approaches such as classification tree, Support Vector Machines, etc., used for intrusion detection. In this paper, thorough investigations have been done on the existing data mining approaches to detect intrusions.. (Abstract)

Keywords- Intrusion Detection System (IDS), intruders, Machine Learning techniques, Data mining

I. INTRODUCTION

Computer networks and their related applications have become an attractive source in the era of information society [1]. Similarly, in recent years, the potential threat to the global information infrastructure has also increased greatly. In order to guard against several cyber attacks and computer viruses, numerous computer security approaches have been extensively researched in the recent years. The major security techniques proposed are cryptography, firewalls, anomaly, intrusion detection, etc. Among the available existing techniques, intrusion detection techniques have been considered to be one of the most significant and competent techniques for protecting complex and dynamic intrusion attacks.

Network intrusion and information safety issues are mainly due to the consequences of extensive internet usage. For example, on February 7th, 2000 the first Denial of Service (DoS) attacks of huge volume were established, aiming the computer systems of huge corporates like Yahoo!, eBay, Amazon, CNN, ZDnet and Dadet [2]. Alternatively, network intrusion is regarded as a new weapon of world war. Thus, it

has become the major concern of the computer society to detect and to prevent intrusions efficiently.

An intrusion is a violation of the security policy of the system, and thus, intrusion detection mainly refers to the methods that detect violations of system security policy. Since the cruelty of attacks in the network has increased radically, Intrusion detection system has become an essential factor to the security infrastructure of several companies. Intrusion detection facilitates companies to defend their systems from various attacks that come with rising network connectivity and dependence on information systems [3].

Recently, intrusion detection techniques through data mining approaches have attracted several researchers. As an essential application area of data mining, intrusion detection focus to lessen the burden of examining vast volumes of audit data and recognizing the performance optimization of detection rules. Several researchers have suggested numerous techniques in various groups, from Bayesian techniques [4] to decision trees [5, 6], from rule based models [7] to functions studying [8]. These techniques have improved the efficiency of the detection to a certain extent.

It is observed from the existing techniques that, most researchers utilized a single algorithm to detect multiple attack classes with miserable performance in certain scenarios. But, detection performance can be greatly improved through complicated technique.

In the present scenario, data mining approaches have taken valuable steps towards solution of several issues in different intrusion detection issues. There are various benefits in utilizing the data mining approaches for solving the problem of network intrusion [9]. Some of the benefits are listed below:

- It can process huge amount of data.
- User's subjective evaluation is not needed, and it is more appropriate to detect the unobserved and hidden information.

Moreover, data mining systems easily performs data summarization and visualization that facilitate the security analysis in various research areas [10].

This paper thoroughly investigates the existing data mining approaches which help in preventing intrusion attacks. The

characteristic features of the intrusion detection techniques are presented in this paper which would facilitate further research in the field of network security.

II. LITERATURE SURVEY

The idea of intrusion detection system was proposed by Anderson in 1980 [11]. Anderson employed statistic technique to examine the behavior of user and to detect those attackers who accesses the system in an unauthorized way. Denning [12] presented a prototype of IDES (Intrusion Detection Expert System) in 1987, then, the concept of intrusion detection system was known progressively, and Denning's approach was considered as a considerable landmark in the area of intrusion detection.

Zenghui and Yingxu [13] proposed a data mining framework for generating intrusion detection models. The main goal is to employ data mining techniques namely, classification, meta-learning, association rules, and frequent episodes to review data for computing misuse and abnormality detection models that correctly capture the actual behavior (i.e., patterns) of intrusions and normal behaviors. Even though, this detection model can significantly detect a considerable percentage of old and new PROBING and U2R attacks, it missed a vast number of new DOS and R2L attacks. Theodoros Lappas and Konstantinos Pelechrinis [14] mostly concentrated on data mining approaches that are being used for dealing with DOS and R2L attacks, and then proposed a new idea on how data mining can help IDSs by utilizing biclustering as a tool to analyze network traffic and improve IDSs.

Sun and Wang [15] presented a new weighted support vector clustering algorithm and utilized it to deal with the problem of anomaly detection. Experimental results reveal the fact that this method obtains high detection rate with low false alarm rate. Su-Yun Wu and Ester Yen [16] compared the performance efficiency of machine learning techniques such as classification tree and support vector machines in intrusion detection system. It is observed from the results that the algorithm of C4.5 for classification tree and SVM are similar to certain level for R2L attack in terms of accuracy, but the accuracy of C4.5 is higher than SVM for other types of attack.

Intruder is one of the most common threats to security. At present, intrusion detection has come out as a significant practice for providing network security. In recent times, data mining approaches have been exploited for the purpose of intrusion detection. The effectiveness of the feature selection techniques is one of the fundamental parameter that has an effect on the success of Intrusion Detection System (IDS). Amudha and Abdul Rauf [17] evaluated the performance of data mining classification approaches specifically, J48, Naive Bayes, NBTree and Random Forest with the use of KDD CUP'99 dataset and mainly concentrated on Correlation Feature Selection (CFS) measure. The results of this evaluation revealed that NBTree and Random Forest performs better than other two approaches based on the predictive

accuracy and detection rate.

Data mining approaches have achieved considerable importance in presenting the helpful information and thereby can assist in improving the decision on recognizing the intrusions (attacks). Panda and Patra [18] evaluated the performance of several rule based classifiers, for instance, JRip, RIDOR, NNge and decision table by using ensemble approach with the intention of constructing an efficient network intrusion detection system. The author exploited KDDCup'99, intrusion detection benchmark dataset (which is a fraction of DARPA evaluation program) for this experimentation. It can be revealed from the outcome that the this scheme is perfect in identifying network intrusions, provides low false positive rate, uncomplicated, consistent and faster in constructing an efficient network intrusion system.

Due to the increase in the number of computer networks at the present scenario, ensuring security in a network against various attacks is essential. Intrusion detection system is one of the popular tools to provide security against the intruders in a network. Exploiting data mining approaches has increased the quality of intrusion detection neither as anomaly detection or misrepresented detection from large scale network traffic operation. Association rule is a popular method to construct quality misused detection. On the other hand, the limitation of association rule is the fact that it often produced with thousands rules which diminishes the performance of IDS. Namik and Othman [19] concentrated on applying post-mining to decrease the number of rules and remaining the most quality rules to generate quality signature. Each partition is mined using Apriori Algorithm, which later carries out post-mining using Chi-Squared (χ^2) computation approaches. The excellence of rules is measured depending on Chi-Square value, which is computed based on the support, confidence and lift of every association rule.

Emerging technologies have metamorphosed the characteristics of surveillance and monitoring application, however the sensory data obtained using different gadgets still remain unreliable and inadequately synchronized. State transition analysis is turning out to be significant components in recognizing intrusions. Ganesh et al., [20] developed a semantic based intrusion detection system in which state transition analysis, pattern matching and data mining techniques are incorporated to enhance the intrusion detection accuracy. Patterns and rules are generated depending on the events identified by WSN. The sink obtains information regarding the numerous actions taking place in the coverage area and correlates the streaming data in spatial domain and time domain. The semantic rules are generated using ANTLR tool.

Networks are safeguarded by means of exploiting several firewalls and encryption software's. However most of these available methods are not adequate and efficient. Majority of the current intrusion detection systems for mobile ad-hoc networks are mostly concentrating on either routing protocols

or only on its effectiveness, but it is unsuccessful to address the security related issues. Some of the nodes which take part in the communication may be selfish, for instance, certain nodes may not forward the packets to the target and by this means it reduces the battery power utilization. In some other cases, certain nodes may act as malicious by initiating security attacks like Denial-of-Service or hack the information. The vital objective of the security solutions for wireless networks is to offer security services, for instance, authentication, confidentiality, integrity, anonymity and availability to mobile users. Esfandi [21] integrates agents and data mining approaches to avoid anomaly intrusion in mobile ad-hoc networks. Home agents present in each system obtain the data from its individual system and by means of data mining approaches the local anomalies are observed. The Mobile agents observe the neighboring nodes and obtain the information from adjacent home agents to find out the correlation between the observed anomalous patterns before it sends the data. This scheme was capable of preventing all the security attacks in an ad-hoc network and reduces the false alarm positive.

Te-Shun Chou and Tsung-Nan Chou [22] proposed a hybrid design for intrusion detection that integrates anomaly detection with misuse detection. This technique also includes an ensemble feature selecting classifier and a data mining classifier. The former includes four classifiers using dissimilar sets of features and each of them utilizes a machine learning algorithm called fuzzy belief k-NN classification algorithm. The latter exploits data mining approaches to automatically obtain computer users' normal behavior from training network traffic data. The outcome of ensemble feature selecting classifier and data mining classifier are then combined together to obtain the final decision.

Several techniques have been developed for intrusion detection using data mining approaches but from the beginning it is uncertain that which data mining approach is most efficient. Zhenwei Yu and Tsai [23] developed a Multi-Class SLIPPER (MC-SLIPPER) scheme for intrusion detection to discover whether there is any significant advantage from boosting dependent learning approach. The fundamental idea is to employ the available binary SLIPPER as a central module, which is a rule learner depending on confidence-rated boosting. Numerous arbitral strategies depending on prediction confidence are developed to judge results from all binary SLIPPER modules.

Security of computers and the networks that connect them is progressively turning out to be much essential. On the other hand, constructing effective intrusion detection techniques with better accuracy and real-time implementation are indispensable. Muntean et al., [24] developed a novel data mining dependent method for intrusion detection by utilizing Cost-sensitive classification together with Support Vector Machines. The author introduced an algorithm that enhances the classification for Support Vector Machines, by multiplying in the training phase the instances of the underrepresented

classes. This technique has exposed that by oversampling the instances of the anomaly and moreover this technique assists the Support Vector Machine algorithm to overcome the soft margin. Consequently, it classifies better future instances of this class of interest.

Some heterogeneous security equipments for instance, firewalls, intrusion detection systems and anti-virus gateways, can generate considerable security events which are complicated to manage effectively. As a result a log-based mining, distributed and multi-protocol supported framework of security monitoring system is developed by Lv Guangjuan et al., [25] and described the structural design of the information security monitoring system. The major concentration is on the correlation analysis engine which illustrates the process that the detection model is constructed using data mining approaches. Security event correlation depending on data mining analysis can automatically obtain association rules, investigate alarming and found new invasion model, and hence it is extremely intelligent technique.

Xin Xu et al., [26] proposed an outline for adaptive intrusion detection with the help of machine learning approaches. Multi-class Support Vector Machines (SVMs) is employed to classifier construction in IDSs and the performance of SVMs is assessed on the KDD99 dataset. Significant results were obtained in the experimental evaluation. For instance, detection rates of 76.7%, 81.2%, 21.4% and 11.2% were obtained for DoS, Probe, U2R, and R2L attacks respectively while False Positive is maintained at the fairly low level of average 0.6% for the four groups. But, this approach can be only employed to a very small set of data (10,000 randomly sampled records) comparing to the huge original dataset (5 million audit records). So, this method is not suitable for all the circumstances and is not regarded as one of the best approach.

Yang Li and Li Guo [27] have already recognized the insufficiency of KDD dataset. However, a supervised network intrusion detection technique depending on Transductive Confidence Machines for K-Nearest Neighbors (TCM-KNN) machine learning algorithm and active learning based training data selection method had been proposed by Yang Li and Li Guo. This new approach was evaluated on a subset of KDD dataset by random sampling 49,402 audit records for the training phase and 12,350 records for the testing phase. An average TP of 99.6% and FP of 0.1% was reported but no additional information about the exact detection rate of each attack categories was presented by the authors..

III. PROBLEMS AND DIRECTIONS

There are various problems and issues present in the existing intrusion detection techniques which are analyzed in this section. This section also provides certain possible solutions to the problems in the existing techniques.

Majority of the intrusion detection techniques available in the literature employed a single algorithm to detect multiple attack categories with miserable performance in most of the

scenarios.

Existing intrusion detection systems are highly dependant on human analysts to distinguish intrusive from non-intrusive network traffic.

Moreover, existing IDSs are developed to detect only particular known service level network attacks. Many attempts have been made to deal with this problem, but resulted in an unacceptable level of false positives. Simultaneously, adequate data exist or could be collected to facilitate network administrators to discover these policy violations. But, the data are so vast and thus, the analysis process takes very long time and the administrators don't have the resources to go through it all and detect the relevant knowledge. Thus, the network administrators don't have the resources to proactively investigate the data for policy violations, particularly in the existence of a high number of false positives that cause them to waste their inadequate resources.

Thus, the most important problem with the existing IDSs approaches is that, the existing IDSs do not provide significant result for all types of attacks.

It is to be understood that, there is considerable variation from one attack category to another and thus, identifying attack category specific algorithm offers a promising research direction for improving intrusion detection performance.

In order to handle the above mentioned problems, an effective and novel research in the areas of data mining and intrusion detection has to be carried out. Efficient machine learning techniques can be used which provide decision aids for the analysts and which automatically generate rules to be used for computer network intrusion detection. Moreover, Neuro-fuzzy techniques can be utilized with better learning techniques to provide precise results in IDS.

IV. CONCLUSION

Intrusion Detection Systems provide the fundamental detection techniques to secure the systems present in the networks that are directly or indirectly connected to the Internet. This paper provides a thorough investigation on the existing intrusion detection techniques through data mining approaches. This paper effectively analysis the problems available in the existing intrusion detection techniques. This paper also suggests certain solutions to the problems available in the existing IDSs. This paper would a suitable platform for the novel researches in the field of network security.

REFERENCES

- [1] Huy Nguyen and Deokjai Choi, "Application of Data Mining to Network Intrusion Detection: Classifier Selection Model", Proceedings of the 11th Asia-Pacific Symposium on Network Operations and Management (APNOMS), Challenges for Next Generation Network Operations and Service Management, Pp. 399-408, 2008.
- [2] Brian Krebs, "A Short History of Computer Viruses and Attacks", <http://www.securityfocus.com/news/2445>.
- [3] L. Vokorokos, A. Kleinova and O. Latka, "Network Security on the Intrusion Detection System Level", Proceedings of International Conference on Intelligent Engineering Systems (INES), Pp. 270-275, 2006.
- [4] G.H. John and P. Langley, "Estimating Continuous Distributions in Bayesian Classifiers", Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence, Pp. 338-345, 1995.
- [5] J. Ross Quinlan, "C4.5: Programs for Machine Learning", Morgan Kaufmann, Publishers Inc. San Francisco, CA, USA, 1993.
- [6] Ron Kohavi, "Scaling up the accuracy of Naïve-Bayes classifier: A decision-tree hybrid", Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Pp. 202-207, 1996.
- [7] Ian H. Witten, Eibe Frank and Mark A. Hall, "Data Mining: Practical Machine Learning Tools and Techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.
- [8] P. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences", PhD Thesis, Harvard University, 1974.
- [9] Ming Xue and Changjun Zhu, "Applied Research on Data Mining Algorithm in Network Intrusion Detection", International Joint Conference on Artificial Intelligence (IJCAI), Pp. 275-277, 2009.
- [10] Eric Bloedorn, Alan D. Christiansen, William Hill, Clement Skorupka, Lisa M. Talbot, Jonathan Tivel, "Data Mining for Network Intrusion Detection: How to Get Started", Technical Paper, 2001.
- [11] J.P. Anderson, "Computer security threat monitoring and surveillance", Technical Report, James P. Anderson Co., Fort Washington, Pennsylvania, 1980.
- [12] D.E. Denning, "An intrusion detection model", IEEE Transaction on Software Engineering, Pp. 222-232, 1987.
- [13] Zenghui Liu and Yingxu Lai, "A Data Mining Framework for Building Intrusion Detection Models Based on IPv6", Proceedings of the 3rd International Conference and Workshops on Advances in Information Security and Assurance, Seoul, Korea, Springer- Verlag, Volume 5576, Pp. 608-618, 2009.
- [14] Theodoros Lappas and Konstantinos Pelechrinis, "Data Mining Techniques for (Network) Intrusion Detection System", 2007.
- [15] Sheng Sun and YuanZhen Wang, "A Weighted Support Vector Clustering Algorithm and its Application in Network Intrusion Detection", First International Workshop on Education Technology and Computer Science (ETCS), Vol. 1, Pp. 352-355, 2009.
- [16] Su-Yun Wu and Ester Yen, "Data mining-based intrusion detectors", Expert Systems with Applications, Vol. 36, No. 3, Pp. 5605-5612, 2009.
- [17] P. Amudha and H. Abdul Rauf, "Performance Analysis of Data Mining Approaches in Intrusion Detection", International Conference on Process Automation, Control and Computing (PACC), Pp. 1-6, 2011.
- [18] M. Panda and M.R. Patra, "Ensembling Rule Based Classifiers for Detecting Network Intrusions", International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom), Pp. 19-22, 2009.
- [19] A.F. Namik and Z.A. Othman, "Reducing network intrusion detection association rules using Chi-Squared pruning technique", 3rd Conference on Data Mining and Optimization (DMO), Pp. 122-127, 2011.
- [20] K.S. Ganesh, M.R. Sekar and V. Vaidehi, "Semantic Intrusion Detection System using pattern matching and state transition analysis", International Conference on Recent Trends in Information Technology (ICRTIT), Pp. 607-612, 2011.
- [21] A. Esfandi, "Efficient anomaly intrusion detection system in adhoc networks by mobile agents", 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), Vol. 7, Pp. 73-77, 2010.
- [22] Te-Shun Chou and Tsung-Nan Chou, "Hybrid Classifier Systems for Intrusion Detection", Seventh Annual Communication Networks and Services Research Conference (CNSR), Pp. 286-291, 2009.
- [23] Zhenwei Yu and J.J.P. Tsai, "A multi-class SLIPPER system for intrusion detection", Proceedings of the 28th Annual International Computer Software and Applications Conference (COMPSAC), Vol. 1, Pp. 212-217, 2004.

- [24] M. Muntean, H. Valean, L. Miclea and A. Incze, "A novel intrusion detection method based on support vector machines", 11th International Symposium on Computational Intelligence and Informatics (CINTI), Pp. 47-52, 2010.
- [25] Lv Guangjuan, Xu Ruzhi, Zu Xiangrong and Deng Liwu, "Information Security Monitoring System Based on Data Mining", Fifth International Conference on Information Assurance and Security, Pp. 472-475, 2009.
- [26] Xin Xu, "Adaptive Intrusion Detection Based on Machine Learning: Feature Extraction, Classifier Construction and Sequential Pattern Prediction", International Journal of Web Services Practices, Vol. 2, No. 1-2, Pp. 49-58, 2006.
- [27] Yang Li and Li Guo, "An Active Learning Based TCM-KNN Algorithm for Supervised Network Intrusion Detection", 26th Computers & Security, Pp. 459-467, 2007.

AUTHORS PROFILE



V. Jaiganesh is working as an Assistant Professor in the Department of Computer Science, Dr. N.G.P. Arts and Science College, Coimbatore and doing Ph.D., in Manonmaniam Sundaranar University, Thirunelveli. He has done his M.Phil., in the area of Data Mining in Periyar University. He has done his post graduate degrees MCA and MBA in Periyar University, Salem. He has presented and published a number of papers in reputed conferences and journals. He has one decade of teaching and research experience and his research interests include Data Mining and Networking.



M. Thenmozhi is working as an Assistant Professor in the Department of Information Technology, Faculty of Engineering, Avinashilingam University for Women, Coimbatore, and doing M.E., Network Engineering in Anna University of Technology, Coimbatore. She received her B.E., at Avinashilingam University for Women, Coimbatore. She has attended various seminars and conferences. She has six years of teaching experience and her interests include Data Mining and Networking.



Dr. P. Sumathi is working as an Assistant Professor in the Department of Computer Science, Chikkanna Government Arts College, Tirupur. She received her Ph.D., in the area of Grid Computing in Bharathiar University. She has done her M.Phil in the area of Software Engineering in Mother Teresa Women's University and received MCA degree at Kongu Engineering College, Perundurai. She has published a number of papers in reputed journals and conferences. She has about fifteen years of teaching and research experience. Her research interests include Data Mining, Grid Computing and Software Engineering.

Trusted On-demand Distance Vector Routing for Ad hoc Networks

Md. Humayun Kabir
Dept. of Computer Science
University of Rajshahi,
Rajshahi, Bangladesh
hkj_cse@yahoo.com

Bimal K. Pramanik
Dept. of Computer Science
University of Rajshahi,
Rajshahi, Bangladesh.
bimal_cst@yahoo.com

Somlail Das
Dept. of Computer Science
University of Rajshahi,
Rajshahi, Bangladesh.
somlail_ru@yahoo.com

Subrata Pramanik
Dept. of Computer Science
University of Rajshahi,
Rajshahi, Bangladesh.
subrata.p@gmail.com

Md. Ekramul Hamid
Dept. of Computer Science
University of Rajshahi,
Rajshahi, Bangladesh
ekram_hamid@yahoo.com

Abstract— This paper presents a new Trusted On-demand Distance Vector (TODV) routing protocol that is dynamic and robust to mitigate the detrimental effects of nodes' malicious behavior, as to provide correct connectivity information. This protocol filters erroneous query and routing information, and determines a route that only involves trustworthy hosts. The operation of TODV is loop free, and can distinguish between local connectivity management (neighborhood detection) and general topology maintenance. When links break, TODV causes the affected set of nodes to be notified so that they are able to invalidate the routes using the lost link. The widely accepted technique in a Mobile Ad hoc Network (MANET) context of route discovery based on broadcasting query packets is the basis of the protocol. The protocol is an enhancement of the Ad hoc On-demand Distance Vector (AODV) routing protocol to ensure that only trustworthy nodes participate in the network. On the other hand it still maintains most of the features of the AODV. The proposed protocol scales to large populations of mobile nodes wishing to form ad hoc networks and can be applied in a wide variety of practical cases.

Keywords- *Ad hoc network, routing protocol, trust, dynamic, broadcasting*

1. INTRODUCTION

Mobile ad hoc networks are self-organizing network architectures in which a collection of mobile nodes with wireless networks interfaces may form a temporary network without the aid of any established infrastructure or centralized administration. According to the IETF definition [1], a mobile ad hoc network is an autonomous system of mobile routers connected by wireless links. This union forms an arbitrary graph. The routers are free to move randomly and organize themselves arbitrarily; thus, the network's wireless topology may change rapidly and unpredictably [2].

Ad hoc networking is a field of very active research in recent years. However, most of the research has been focused around various protocols for multi-hop routing, leaving the area of security mostly unexplored. At the same time, new applications of ad hoc networking, including wireless sensor networks, ubiquitous computing and peer-to-peer applications,

introduce a need for strong privacy protection and security mechanisms.

High level security requirements for ad hoc networks are basically identical to security requirements for any other communications system, and include following services:

- authentication
- confidentiality
- integrity
- non-repudiation
- access control
- availability

However, similar to wireless communication systems creating additional challenges for implementation of aforementioned services when compared to fixed networks, ad hoc networks can be viewed as even more extreme case, requiring even more sophisticated, efficient and well designed security mechanisms [3][4][5][6]. These additional challenges are caused by two basic assumptions of an ad hoc system:

1. lack of the infrastructure, and
2. a very dynamic and ephemeral character of the relationships between the network nodes.

The lack of infrastructure implies that there is no central authority, which can be referenced when it comes to making trust decisions about other parties in the network and that accountability cannot be easily implemented. The transient relationships do not help in building trust based on direct reciprocity and give additional incentives to nodes to cheat.

Ad hoc networks rely on cooperation of involved nodes in order for the network to emerge and operate. Current versions of mature ad hoc routing algorithms only detect if the receiver's network interface is accepting packets, but they otherwise assume that routing nodes do not misbehave. Whereas such an assumption may be justified where single domains are concerned, it is not easy to transpose it on a network consisting of nodes, unknown to, and untrusted by, each other. Since ad hoc networks deploy multi-hop routing

protocols, where each of the nodes in addition to its own packets has to forward packets belonging to other nodes, selfish behavior may represent a significant advantage for a node, saving his battery power and reserving more bandwidth for its own traffic. However, if a large number of nodes start to behave non-cooperatively, the network may break down completely, depriving all users of services. Non-cooperative behavior in multi-hop routing protocols may also result in a denial of service attacks on the network, where the malicious nodes join the network for a sole reason of misbehaving and depriving all other nodes of legitimate services. Such denial-of-service focused misbehavior may consist of dropping (not forwarding) the packets, injecting incorrect routing information, replayed expired routing information or distorting routing information in order to partition the network [7]. Also bogus nodes may try to attract as much traffic as possible to themselves in order to be able to analyze it. In general, attacks on a routing protocol can be classified as [8]:

- non forwarding
- traffic deviations and route modifications
- lack of error messages
- frequent route updates.

Finding efficient solution to these problems in an open ad hoc environment is still an open issue.

In ad hoc networks, it is hard to employ static routes; link-state based routing protocols and complex public-key encryption algorithms. Routing protocols must be dynamic and robust against malicious attacks [9]. Our proposed Scheme, Trusted On-Demand Distance Vector Routing (TODV), is an enhancement of the AODV routing protocol [10], which maintains most of the advantages of the AODV routing protocol. Until now Most of the proposed secure MANET routing protocols [5][7][11][12][13][14][15][16] assumed some kind of a priori secret association or key exchange between the nodes, while our proposed scheme does not make use of such an assumption. An effort return based trust model proposed by Pirzada et al. in [17] for pure ad hoc networks requires the participating nodes in AODV routing protocol to support the features such as promiscuous mode operation, omni directional transceivers, but it requires complex calculations to establish normalized events . A simple trust model based on packet forwarding ratio to evaluate neighbours' behaviours is proposed in by Xin et al. in [18]. The author proposed a multipath reactive routing protocol (AOTDV) to discover trustworthy forward paths and alleviate the attacks of malicious nodes to meet the dependable or trust requirements of data packets. This model also requires complex calculations.

Our proposed scheme discovers a fully trusted path between the source and destination that consists of only trusted nodes. The widely accepted technique in the MANET context of route discovery based on broadcasting query packets is the basis of our protocol. The broadcast nature of the radio signals mandates that each transmission is received by all the neighbors, which are assumed to operate in promiscuous mode, that is, able to overhear all transmissions from nodes within the range of their transceiver. Nodes operating in promiscuous

mode overhear the transmissions of their successors and may verify whether the packet was forwarded to the downstream node and check the integrity of the forwarded packet. When links break, TODV causes the affected set of nodes to be notified so that they are able to invalidate the routes using the lost link.

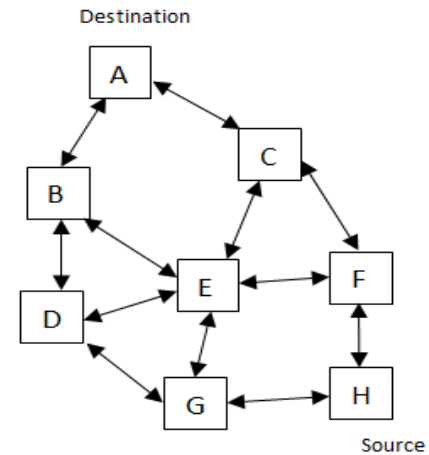


Fig-1: Nodes operating in promiscuous mode

2. AN INTRODUCTION: AODV ROUTING PROTOCOL

The Ad hoc On Demand Distance Vector (AODV) routing algorithm is a routing protocol designed for a Mobile Ad hoc Network (MANET). AODV is capable of both unicast and multicast routing and an On demand algorithm that builds routes between nodes only as desired by source nodes. Here the routes are created and maintained only when they are needed. For that a routing table stores the information about the next hop to the destination and a sequence number indicating the freshness of the received information. New version of the AODV routing protocol [10] has also a feature that only the destination host can reply to the sent request. When the reply is sent back to the requested host the actual hop metric is counted. The intermediate hosts records information about the replied host upon receiving the reply message. The hosts must record and forward new information only when the sequence number is greater or if the sequence number is the same and hop metric is smaller. When a node wants to communicate with a destination while it obtain no proper route entry for that destination, the source node will broadcast an RREQ (Routing REQuest) message to all its neighbors. Each neighbor who receives this RREQ will check in its own routing table.

If not contains route entry: set up a reverse path towards the originator of RREQ and rebroadcast this routing request.

If contains route entry: will generate an RREP (Routing REPLY) message and unicast it to the next hop toward the originator of the RREQ, as indicated by the routing entry for that originator. When a node receives an RREP message, it

first updates some fields of the route table and the routing reply, and then forwards it to the next hop towards the originator. In this way, this RREP will ultimately reach the source node and setup a path for two way communication.

2.1 The Proposed: TODV Routing Protocol

The main goal in the proposed TODV protocol is to establish a trusted route path between the source and destination, so as to avoid any kind of directed attacks. In fact, most of the routing disruption attacks are caused by malicious injection or altering of routing data. So, we feel that there is a need to prevent these attacks by totally hiding the routing information from unauthorized nodes.

2.1.1 Assumptions

In this work, we make some assumptions and establish the trusted route on demand basis from source to destination. Although TODV does not depend specifically on particular aspects of the physical medium across which packets are disseminated, its development has been largely motivated by limited range broadcast media such as those utilized by infrared or radio frequency wireless communication adapters. Using such media, a mobile node can have neighbors, which hear its broadcasts and yet do not detect each other (the hidden terminal problem). No attempt is made to use specific characteristics of the physical medium in the proposed system, nor to handle specific problems posed by channelization needs of radio frequency transmitters. Nodes that need to operate over multiple channels are presumed to be able to do so. The only requirement placed on the broadcast medium is that neighboring nodes can detect each other's broadcasts, which are assumed to operate in promiscuous mode, that is, able to overhear all transmissions from nodes within the range of their transceiver. It is assumed that TODV uses symmetric links between neighboring nodes.

2.1.2 TODV: An Overview

Our proposed Scheme, Trusted On-Demand Distance Vector Routing (TODV), enables dynamic, self-starting, multi-hop routing among participating mobile nodes wishing to establish and maintain an ad hoc network. TODV allows mobile nodes to obtain trusted routes quickly for new destinations, and does not require nodes to maintain routes to destinations that are not in active communication. TODV also defines timely responses to link breakages. The operation of TODV is loop free, and can distinguish between local connectivity management (neighborhood detection) and general topology maintenance. When links break, TODV causes the affected set of nodes to be notified so that they are able to invalidate the routes using the lost link.

The widely accepted technique in the MANET context of route discovery based on broadcasting query packets is the basis of this protocol. The broadcast nature of the radio signals

mandates that each transmission is received by all the neighbors, which are assumed to operate in promiscuous mode, that is, able to overhear all transmissions from nodes within the range of their transceiver. Nodes operating in promiscuous mode overhear the transmissions of their successors and may verify whether the packet was forwarded to the downstream node and check the integrity of the forwarded packet. Upon detection of a misbehaving node for a route discovery packet, the predecessor node enters the identity of the misbehaving node with the identification of the route discovery packet for which the misbehaving node misbehaves into its black list. This information is maintained for at least enough time for the route discovery packet to traverse the network and produce a reply to the sender. The node tags this information so that in later time it can identify any reply coming from the misbehaving node for that route discovery packet and not processes or unicast the reply.

Route Requests (RREQs) and Route Replies (RREPs) are the two message types defined by the proposed scheme. As long as the endpoints of communication connection have valid routes to each other, the proposed protocol does not play any role. When a route to a new destination is needed, the node uses a broadcast RREQ to find a route to the destination. A route can be determined when the request reaches the destination itself. The route is made available by unicasting a RREP back to the source of the RREQ. Since each node receiving the request keeps track of a route back to the source of the request, the RREP can be unicast back from the destination to the source.

If a RREP is broadcast to the limited broadcast address, the time-to-live (TTL) value of one, the destination sequence number as the latest destination sequence number and a destination address of the node's address itself then it is received by all the node's neighbors, and treated by them as a "hello" message. This hello message is a local advertisement for the continued presence of the node. Neighbors that are using routes through the broadcasting node will continue to mark the routes as valid. If hello messages from a particular node stop coming, the neighbor can assume that the node has moved away or down. When that happens, the neighbor will mark the link to the node as broken, and may trigger a notification to its active neighbors that the link has broken. A neighbor is considered active for that destination if it originates or relays at least one packet for that destination within the most recent `active_route_timeout` period.

The proposed routing protocol deals with routing table management. Routing information is kept for all known routes and it uses the following fields with each routing table entry: destination address, next hop address, lifetime (expiration or deletion time of the route), hop Count (number of hops to reach the destination), active Neighbors for that route and the destination sequence number from the RREP packet.

2.1.3 Detailed Protocol Description

This section describes the scenarios under which nodes generate and forward RREQ and RREP packets, and how the fields in the packets are handled. In this section Nodes also detects Misbehavior of their neighbors by checking RREQ and RREP packets, and decides whether or not to forward RREQ or RREP packets.

2.1.3.1 Generating Route Requests

A node disseminates a RREQ packet when it determines that it needs a route to a destination and does not have one available. This can happen if the destination is previously unknown to the node, or if a previously valid route to the destination expires or is marked as invalid. The RREQ contains the following fields:

`<source_addr, broadcast_id, dest_addr, hop_cnt, rcv_node_id, snd_node_id >`

The broadcast_id field is incremented by one from the last broadcast_id used by the current node. Each node maintains only one broadcast_id. The hop_cnt field is set to zero. The rcv_node_id is set to a null value. The snd_node_id is set to the ID of the originating node.

Before broadcasting the RREQ packet, the originating node buffers the information of the RREQ packet into its history table. In this way, when the node receives the packet again from its neighbors, it will not re-forward the packet.

After broadcasting a RREQ packet a node waits for a RREP, and if the reply is not received within a pre-established time (in milliseconds), the node may rebroadcast a new RREQ packet. The RREQ packet may be rebroadcast up to a maximum number of times (pre-established). Each rebroadcast has to increment the broadcast_id field.

2.1.3.2 Misbehave Detection by Checking RREQ Packet And Processing and Forwarding Route Requests

When a node receives a broadcast RREQ packet, the node first checks its history table to see whether the node has received a RREQ packet before with the same source_addr and broadcast_id fields. If such a RREQ packet has been received before, the node verifies the rcv_node_id field of the newly received RREQ packet to it's own ID. If the node finds a match between the rcv_node_id field of the newly received RREQ packet and it's own ID, the node then verifies the various fields of the newly received RREQ packet with the fields of the RREQ packet buffered into its history table. If any mismatch is found, the node records the ID of the misbehaving node from which the new RREQ packet is received (which is obtained by the snd_node_id field from the newly received RREQ packet) into its black list. The node also records the source_addr and broadcast_id fields of the newly

received RREQ packet into its black list to indicate for which RREQ packet the misbehaving node misbehaves, and then simply discards the RREQ packet. These information are used at later time not to relay or process any RREP packet for the tamping RREQ packet that are coming from the misbehaving node. These information are maintained for at least enough time for the RREQ packet to traverse the network and produce a reply to the sender. If no mismatch is found the node silently discards the newly received RREQ packet.

If the node finds no match between the rcv_node_id field of the newly received RREQ packet and it's own ID, the node does not processes the newly received RREQ packet further and silently discards the newly received RREQ packet.

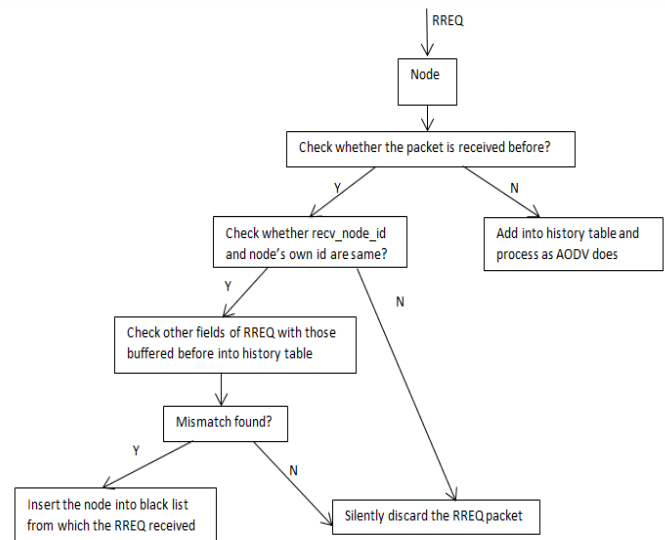


Fig.-2: Pictorial view of misbehave detection by checking RREQ packet

If no such RREQ packet is found in its history table, the node first increments the hop_cnt field value of the newly received RREQ packet, and then buffers the fields of the received RREQ packet into its history table. The node also stores the following information from the received RREQ packet into its reverse list in order to implement the reverse path setup that will accompany the transmission of the eventual RREP:

- source_addr
- broadcast_id
- dest_addr and
- snd_node_id

These reverse path route entries are also maintained for at least enough time for the RREQ packet to traverse the network and produce a reply to the sender. The node then first sets the rcv_node_id field by the snd_node_id field and then the snd_node_id field by its own ID of the received RREQ packet. Finally, the node rebroadcasts the received RREQ packet with the same values in the other fields.

2.1.3.3 Generating Route Replies

Upon reception of a RREQ packet, a node must generate a RREP packet if it is the destination. The RREP packet contains the following fields:

`<source_addr, broadcast_id, dest_addr, hop_cnt, dest_seq_no, snd_node_id, lifetime>`

When generating a RREP packet, a node copies the source_addr, broadcast_id and dest_addr from the received RREQ packet into the corresponding fields of the RREP packet. The destination node places its own id into the snd_node_id field of the RREP packet, and sets the value zero to the hop_cnt field of the RREP packet. The dest_seq_no is set to the sequence number associated with the destination node. The destination node also sets the lifetime field of the RREP packet by a time value for which nodes receiving the RREP packet consider the route to be valid. Once created, the RREP packet is unicast to the next hop toward the originator of the RREQ packet, indicated by the snd_node_id of the last received RREQ packet for which the RREP packet is generating.

2.1.3.4 Misbehave Detection by Checking RREP Packet And Processing and Forwarding Route Replies

When a node receives a RREP packet, the node first checks the hop_cnt field value in the RREP packet to know whether it is zero. If hop_cnt field value is zero the node then checks the following conditions (To check whether the successor node replies truly):

1. At first the node checks the dest_addr in the RREP packet with the dest_addr recorded in the history table for the same <source_addr, broadcast_id> pair. If not equal, does not process the RREP packet further (i.e., simply drops the RREP packet). If equal, checks the next condition.
2. The node checks the dest_addr and snd_node_id fields in the RREP packet. If not equal, does not process the RREP packet further (i.e., simply drops the RREP packet). If equal, the node processes the RREP packet according to the following conditions.

The node finds out a match between the snd_node_id of the RREP packet with the snd_node_id of the RREQ packet buffered into the black list for which the RREP packet is. If found the node simply drops the RREP packet.

If the node does not in the black list from which the RREP came, the node increments the hop_cnt value in the RREP packet by one, to account for the new hop through the intermediate node. Call this incremented value the "New Hop Count". Then the forward route for this destination is created if it does not already exist. Otherwise, the node compares the

dest_seq_no in the RREP packet with its own stored destination sequence number for the Destination in the RREP packet. Upon comparison, the existing entry is updated only in the following circumstances:

1. the dest_seq_no in the RREP is greater than the node's copy of the destination sequence number, or
2. the sequence numbers are the same, but the route is marked as inactive, or
3. the sequence numbers are the same, and the New Hop Count is smaller than the hop count in route table entry.

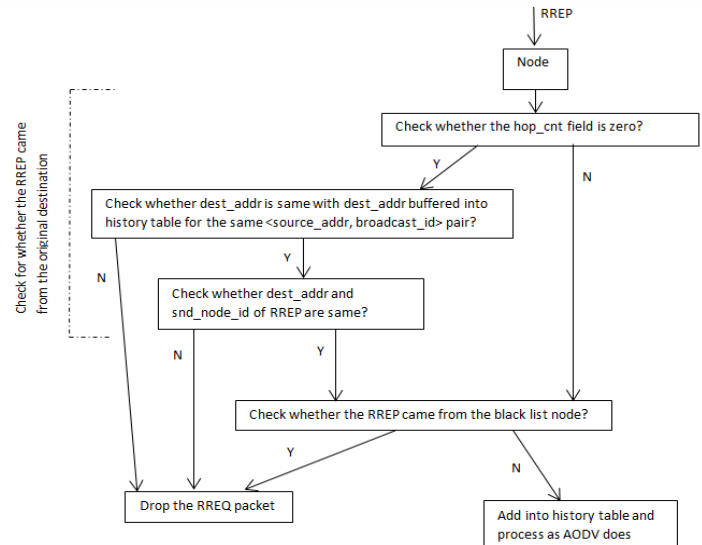


Fig.-3: Pictorial view of misbehave detection by checking RREP

If the route table entry to the destination is created or updated, then the following actions occur:

1. the route is marked as active,
2. the next hop in the route table entry is assigned to be the node from which the RREP packet is received, which is obtained from the snd_node_id field of the RREP packet,
3. the hop count is set to the value of the New Hop Count,
4. the expiry time is set to the current time plus the value of the lifetime in the RREP packet.
5. and the destination sequence number is the dest_seq_no in the RREP packet.

The current node can subsequently use this route to forward data packets to the destination.

If the current node is the node indicated by the source_addr in the RREP packet AND a forward route has been created or updated as described above, then route discovering is successful.

If the current node is not the node indicated by the source_addr in the RREP packet AND a forward route has been created or updated as described above, the node consults its reverse list entry for the originating node to determine the next hop for the RREP packet. The node places its own ID into the snd_node_id field of the RREP packet, and then forwards the RREP packet towards the originator using the information in that reverse list entry.

2.1.3.5 Generating Hello Messages

Every node generates a "hello" message once every HELLO_INTERVAL milliseconds. This hello message is a broadcast RREP with TTL = 1, and the message fields set as follows:

Destination Address

The node's address

Destination Sequence Number

The latest sequence number

Hop Count

0

Lifetime

$(1 + \text{ALLOWED_HELLO_LOSS}) * \text{HELLO_INTERVAL}$

2.1.3.6 Initiating Triggered Route Replies

A node can trigger an unsolicited RREP if either it detects a link breakage for a next hop along an active route in its route table, or if it receives a RREP from a neighbor with an infinite metric for an active route (i.e., containing a Destination Address for which there is a route table entry with a nonempty active-list).

The unsolicited RREP is unicast to each neighbor in the nonempty active-list for the route to that destination. The contents of the RREP fields are set as follows:

Hop Count

A large number

Destination Address

The destination in the broken route

Destination Sequence Number

One plus the destination sequence number recorded in the route.

2.1.3.7 Detecting Link Breakage

A node can detect a link breakage by listening to "hello" messages from its neighbors. If it has received hello messages from a particular neighbor, but misses more than ALLOWED_HELLO_LOSS consecutive hello messages from that neighbor, the node can presume that the particular neighbor is no longer able to maintain a direct link with the mobile node. When this happens, the node should assume that

its link with the former neighbor has been broken, and proceed as in Section 2.3.6. A node should assume that a hello message has been missed if it is not received within double times the duration of the HELLO_INTERVAL.

Alternatively, the node can use any physical-layer or link-layer methods to detect link breakages with nodes it has considered as neighbors.

3. DISCUSSION

An interesting characteristic of the proposed routing protocol is that it does not make use of a priori secret association or key exchange between the nodes. The proposed scheme discovers a fully trusted path between the source and destination that consists of only trusted nodes. In the following, an example of a snapshot of the network is described in which a problem may arise.

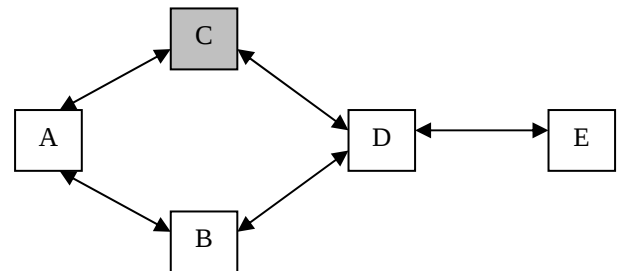


Fig. 4: A snapshot of a network in which a problem may arise.

As in the above figure 4, it is assumed that node A wants to send data to node E. So, node A broadcast a RREQ packet requesting to set up a route to node E. Also it is assumed that here node C is a malicious node. In this snapshot, when node D hears the broadcast of node B before node C, there is no problem, that is, a route is established between node A and node E as A-B-D-E. However, when node D hears the broadcast of node C before node B, there is a problem that no route is established between node A and node E, although a route is available between node A and E as A-B-D-E. In this case, when node D hears the broadcast of node B it drops the RREQ packet due to duplicates. So, no route is established between node A and E. It is being worked on to solve this problem.

To relieve from IP spoofing (Any intermediate node may hide its real IP address or MAC address and uses different one) the proposed protocol may be able to provide correct and current connectivity information. Each node in the network may maintain a neighbor list to determine whether or not a node is its neighbor from which a message has come. The proposed protocol may maintain the neighbor list by hearing the consecutive hello messages.

As the next step of the research, it will be tried to present a detailed performance evaluation of the proposed TODV routing protocol for various network instances and node processing

capabilities. It will also be tried to evaluate the overhead of the protocol with respect to existing protocols, in normal, non-faulty conditions as well as in adversarial environment.

CONCLUSIONS

In this paper an efficient trusted routing protocol for mobile ad hoc networks has been proposed that guarantees the discovery of fresh and correct connectivity information over an unknown network, in the presence of malicious nodes. The protocol introduces a set of features, such as the requirement that when the route discovery packet arrives at the destination, the destination node replies, and only those replies arrive at the destination that not relaying through the misbehaving nodes over the reverse route of the route discovery packet, the acceptance of route error messages only when generated by nodes on the active route, the protection of the looping due to continuous broadcasting of the duplicate route discovery packet and the regulation of the route discovery packet propagation.

The resultant protocol is capable of operating without the existence of a priori secret association or key exchange between the nodes. Its sole requirement is the widely accepted technique in the MANET context of route discovery based on broadcasting query packets. The broadcast nature of the radio waves mandates that each transmission is received by all neighbors, which are assumed to operate in promiscuous mode (i.e., able to overhear all transmissions from nodes within the range of their transceiver). Nodes operating in promiscuous mode overhear the transmissions of their successors and may verify whether the packet was forwarded to the downstream node and check the integrity of the forwarded packet.

Simulating a routing protocol is a crucial step in verifying the correct design and operation of the protocol. A simulation of the TODV routing protocol is the next step of this research. As TODV continues to be refined, it is possible that further changes will be required. We look forward to the completion of the implementations, the design of a test bed in which to test the implementation, and interoperability testing with other existing methods.

REFERENCES

- [1] B. Shrader, "A proposed definition of Ad hoc network," Royal Institute of Technology (KTH), Stockholm, Sweden, May 2002.
- [2] M. M. Lehmus, "Requirements of Ad hoc Network Protocols," Technical report, Electrical Engineering, Helsinki University of Technology, May 2000.
- [3] Jean-Pierre Hubaux, Levente Buttyan, and Srdan Capkan, The quest for security in mobile ad hoc networks, Proceedings of the ACM symposium on Mobile Ad hoc Networking and Computing (MobiHOC) (Long Beach, CA), ACM, Oct. 2001.
- [4] Adrian Perrig, Robert Szewczyk, Victor Wen, David Culler, and J. D. Tygar, SPINS: Security protocols for sensor networks, 7th ACM International Conference on Mobile Computing and Networking (Rome, Italy), vol. 1, ACM Press, 2001, pp. 189-199.
- [5] Seung Yi, Prasad Naldurg, and Robin Kravets, Security-aware ad hoc routing for wireless networks, Technical Report, USA, Aug. 2001.
- [6] Lidong Zhou and Zygmunt Hass, Securing ad hoc networks, IEEE network magazine 13, (1999), no. 6.

- [7] Sergio Merti, T. J. Giuli, Kevin Lai, and Mary Baker, Mitigating routing misbehavior in mobile ad hoc networks, Mobile Computing and Networking, 2000, pp. 255-265.
- [8] Sonja Buchegger and Jean-Yves Le Boudec, Performance analysis of the CONFIDENT protocol (Cooperation of nodes: Fairness In dynamic Ad hoc Networks), Proceedings of MOBIHOC'02 (EPFL Lausanne, Switzerland), ACM, June 9-11, 2002.
- [9] Bharat Bhargava, Trusted Route Discovery in Ad hoc Networks, CERIAS and Department of Computer Science, Purdue University, West Lafayette, IN 47907-1398, USA.
- [10] C. Perkins, E. Belding-Royer and S. Das "Ad hoc On-Demand Distance Vector (AODV) Routing", RFC 3561, IETF Network Working Group, July 2003.
- [11] P. Papadimitratos, "Secure Routing: Methods for Protecting Routing Infrastructures – A Survey," work in progress.
- [12] L. Buttyan and J.P. Hubaux, "Enforcing Service Availability in Mobile Ad Hoc WAnS," 1st MobiHoc, BA Massachusetts, Aug. 2000.
- [13] M. Guerrero, "Secure AODV", Internet draft sent to manet@itd.nrl.navy.mil mailing list, Aug. 2001.
- [14] P. Papadimitratos and Z.J. Haas, "Secure Message Transmission in Mobile Ad Hoc Networks," submitted for publication.
- [15] P. Papadimitratos and Z. Haas. Secure routing for mobile ad hoc networks. In Proc. SCS Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS), 2002.
- [16] Alec Yasinsac and Stephen Carter, Secure Position Aided Ad hoc Routing, Florida State University, 2002.
- [17] Pirzada et al., "Performance Comparison Of Trust-Based Reactive Routing Protocols", IEEE Transactions on Mobile Computing, Vol. 5, No. 6, June 2006, pp. 695-710.
- [18] Xin Li, Zhiping Jia, Peng Zhang, Ruihua Zhang and Haiyang Wang, "Trust-based On-demand Multipath Routing in Mobile Ad Hoc Networks", IET Information Security. 2010, 4. (ISSN: 1751-8709).

AUTHORS PROFILE



Md. Humayun Kabir is currently working as Lecturer in the Department of Computer Science & Engineering, University of Rajshahi, Bangladesh. He received B.Sc. (Hons.) and M.Sc. degree in Computer Science & Engineering from the same university. In addition he is currently pursuing the MPhil. degree with the Department of Computer Science & Engineering, University of Rajshahi, Bangladesh. His research interests include Network Security, Internet and mobile computing, Mobile Ad hoc Networks and wireless sensor networks.



Bimal K. Pramanik received the B.Sc. degree from the University of Rajshahi, Rajshahi, Bangladesh, in 1996 and the M.Sc. degree from the department of Microelectronics and Information Technology, Royal Institute of Technology, Stockholm, Sweden, in 2004. He has received his Ph.D. from the department of Electronic and Phonic Systems Engineering, Kochi University of Technology, Japan. Now he is working as Associate Professor in The department of Computer Science & Engineering, University of Rajshahi, Bangladesh.



Subrata Pramanik did his M.Sc. in Computer Science and Technology from University of Rajshahi, Bangladesh and again in Computer Science from University of Northern British Columbia, Canada. Currently he has been working as an Associate Professor in Dept. of Computer Science & Engineering, University of Rajshahi, Bangladesh.



Somlail Das received his B.Sc and M.Sc degree from the Department of Applied Physics and Electronics, Rajshahi University, Bangladesh. During 1998-2001, he was a lecturer in the Department of Computer Science and Engineering, Rajshahi University. He is currently working as an Associate Professor in the same Department. His research interests include Digital Signal Processing and Speech Enhancement.



Md. Ekramul Hamid received his B.Sc and M.Sc degree from the Department of Applied Physics and Electronics, Rajshahi University, Bangladesh. After that he obtained the Masters of Computer Science from Pune University, India. He received his PhD degree from Shizuoka University, Japan. During 1997-2000, he was a lecturer in the Department of Computer Science and Engineering, Rajshahi University. Since 2007, he has been serving as an Associate Professor in the same department. He was an assistant professor in the college of computer science at King Khalid University, Abha, KSA from 2009-2011. His research interests include Digital Signal Processing and Speech Enhancement.

An Intrusion Detection System Framework for Ad Hoc Network

Arjun Singh 1

Dept. of Computer Science & Engineering
Sir Padampat Singhania University
Udaipur, India
arjun.singh@spsu.ac.in

Surbhi Chauhan 2

Dept. of Computer Science & Engineering
Amity University
Noida, India
Surbhichauhan2009@gmail.com

Kamal Kant 3

Dept. of Computer Science & Engineering
Amity University
Noida, India
kamalkant25@gmail.com

Reshma Doknaia 4

Sr. Software Engineer
BMC Pvt. Ltd.
Pune, India
reshma.dokania@gmail.com

Abstract— Secure and efficient communication among a set of mobile nodes is one of the most important aspects in ad-hoc wireless networks. Wireless networks are particularly vulnerable to intrusion, as they operate in open medium, and use cooperative strategies for network communications. By efficiently merging audit data from multiple network sensors, we analyze the entire ad hoc wireless network for intrusions and try to inhibit intrusion attempts. This paper presents an intrusion detection system for ad hoc network, which uses reputation system to minimize the usage of battery power and bandwidth.

Keywords-IDS, LID, MDM,ADM,SSD

I. INTRODUCTION

Ad hoc network are dynamic, peer-to-peer networks that do not have a pre-existing infrastructure and are characterized by wireless multi-hop communication. The unreliability of wireless links between nodes, constantly changing topology due to the movement of nodes in and out of the network, and lack of incorporation of security features in statically configured wireless routing protocols not meant for ad hoc environments all lead to Increased vulnerability and exposure to attacks. Securing wireless ad hoc networks is particularly difficult for many reasons including the following:

- **Vulnerability of channels.** As in any wireless network, messages can be eavesdropped and fake messages can be injected into the network without the difficulty of having physical access to network components.
- **Vulnerability of nodes.** Since the network nodes usually do not reside in physically protected places, such as locked rooms, they can more easily be captured and fall under the control of an attacker.
- **Absence of infrastructure.** Ad hoc networks are supposed to operate independently of any fixed infrastructure. This makes the classical security solutions based on certification authorities and on-line servers inapplicable.

- **Dynamically changing topology.** In mobile ad hoc networks, the permanent changes of topology require sophisticated routing protocols, the security of which is an additional challenge. A particular difficulty is that incorrect routing information can be generated by compromised nodes or as a result of some topology changes and it is hard to distinguish between the two cases.

II. EINTRUSION DETECTION IN WIRELESS AD HOC NETWORK

Intrusion Detection Systems (IDS) may be classified based on the data collection maintaining the integrity of the specifications mechanism, as well as the technique used to detect events. While the requirement of intrusion detection for both fixed wired and wireless ad-hoc networks are the same, wireless ad-hoc networks impose additional challenges. The effectiveness of IDS solutions that were designed for fixed wired networks is limited for wireless ad-hoc network, as described below:

- Wireless ad-hoc networks lack key concentration points where network traffic can be monitored. This limits the effectiveness of a network-based IDS sensor, since only the traffic generated within radio transmission range may be monitored.
- In a dynamically changing ad-hoc network, it may be difficult to rely on the existence of a centralized server to perform analysis and correlation.
- The secure distribution of signatures may be difficult, due to the properties of wireless communication and mobile nodes that operate in disconnect mode.

Intrusion detection can be classified into three broad categories:

1. Anomaly detection, signature
2. Misuse detection, and
3. Specification based detection.

A. Anomaly Detection

In an anomaly detection system a baseline profile of normal system activity is created. Any system activity that deviates from the baseline is treated as a possible intrusion.

The problems with strict anomaly detection are that:

- Anomalous activities that are not intrusive are flagged as intrusive.
- Intrusive activities that are not anomalous result in false negatives.

One disadvantage of anomaly detection for mobile computing is that the normal profile must be periodically updated and the deviations from the normal profile computed. The periodic calculations can impose a heavy load on some resource constrained mobile devices; perhaps a lightweight approach that involves comparatively less computation might be better suited.

B. Misuse Detection

In misuse detection, decisions are made on the basis of knowledge of a model of the intrusive process and what traces it ought to leave in the observed system. *Legal* or *illegal* behavior can be defined and observed behavior compared accordingly. Such a system tries to detect evidence of intrusive activity irrespective of any knowledge regarding the background traffic (i.e., the normal behavior of the system).

C. Specification-Based Detection

This defines a set of constraints that describe the correct operation of a program or protocol, and monitors the execution of the program with respect to the defined constraints. This technique may provide the capability to detect previously unknown attacks, while exhibiting a low false positive rate.

III. INTRUSION DETECTION ARCHITECTURE

Each node on the ad hoc network has an IDS agent running on it. The IDS agents work together through cooperative intrusion detection to decide when and how the network is being attacked. The architecture is divided into two parts: the mobile IDS agent, which resides on each node in the network, and the stationary secure database, which contains global signatures of known misuse attacks and stores patterns of each user's normal activity in a non-trusted environment. An IDS agent runs at each mobile node does local intrusion detection independently, and neighboring nodes collaboratively work on a larger scale. Individual IDS agents placed on each and every node run independently

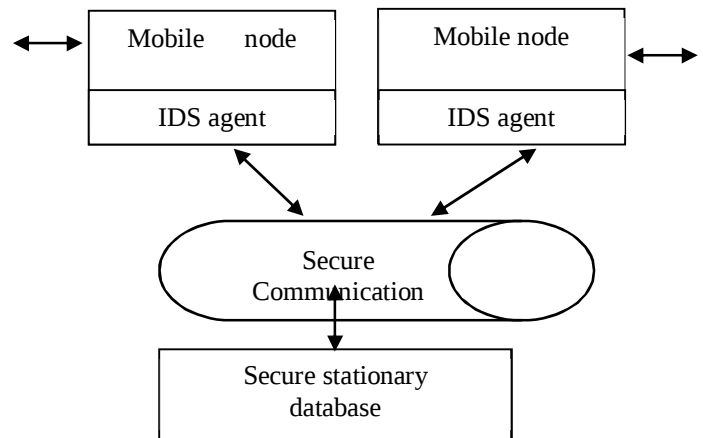


Figure1. Architecture of IDS

and monitor local activities, detect intrusions from local traces, and initiate responses.

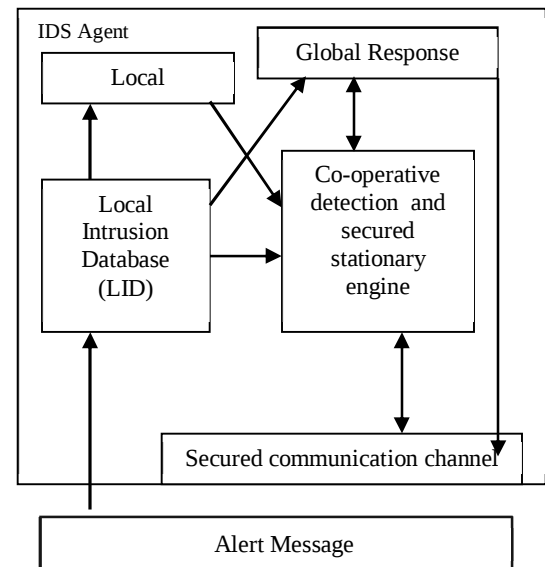


Figure 2. IDS Agent Architecture

Neighboring IDS agents cooperatively participate in global intrusion detection actions when an anomaly is detected in local data. The data collection module gathers local audit traces and activity logs that are used by the local detection engine to detect local anomaly. Detection methods that need broader data sets or require collaborations among local IDS agents use the cooperative detection engine. Both the local and global response modules provide intrusion response actions. The local response module triggers actions local to this mobile node, while the global one coordinates actions among neighboring nodes, such as the IDS agents in the network electing a suitable action. A secure communication module provides a high-confidence communication channel among IDS agents.

IV. REPUTATION MECHANISM

Reputation mechanism is used within ad hoc networks to address some of the threats arising from misbehaving network nodes. These mechanisms are potentially of particular value in addressing the threats arising from selfish nodes. In the context of ad hoc networks, these mechanisms seek to dynamically assess the trustworthiness of neighboring network nodes, with a view to excluding untrustworthy nodes. There are three types of reputation, which are combined to form a global reputation value for a community member. Each calculation is normalized so that a reputation values ranges from -1(bad) to +1 (good). 0 represents a neutral view, and this is used when there is not enough observation to make an accurate assessment of a node's reputation. The three reputation types are as follows:

1. Subjective reputation is locally calculated, where node A calculates the reputation of a neighbor node B at a given time for a particular function.
2. Indirect reputation are accepted by node A from node C about node B. only positive reputation values are used, to eliminate an attack where a malicious node transmits negative reputation information to cause a denial-of-service.
3. Functional reputations are related to a certain function where each function is a weight as to its importance. Each node maintains a reputation table. This table contains of the reputations of other nodes, with each entry consisting of a unique ID, recent subjective observation, recent indirect observations and the composite reputation for a given function. Thus a reputation table has to be maintained for each function that is to be monitored.

There are 3 ways in which a reputation table is updated.

1. A node A requests a service from node B, but node B refuses to perform the service. Thus node A will decrease its perceived reputation of node B. this is a calculation of node B's subjective reputation.
2. A global distribution of reputation takes place within a reputation dissemination phase. This phase involves sending messages containing a list of entities, which have successfully co-operated in providing a function, i.e., a list of nodes with positive reputation.
3. The reputation is gradually decreased to a null value if there is no interaction with observed node.

When a node A, with a good reputation, is asked to perform a service by a node B, who has a bad reputation Node A can refuse to cooperate in doing so. Node A is required to send a message to all nodes in the ad hoc network, stating that it is denying services to node B. The neighbor nodes of A and B must check that node B's reputation is negative in their own reputation table. If one of the neighbor nodes does not agree with node A's negative reputation value for node B, then this neighbor node decreases

the reputation of node A, i.e., the node which sent the denial of service message.

Reputation system alerts path manager. The path manager ranks routed according to security metric. All paths, which contain a bad behaving node, are deleted. The path manager also decides what to do with requests received from badly behaved nodes. The local intrusion detection system (LIDS) is distributed in nature and utilizes mobile agents on each of the nodes of the ad hoc network. In order to make local intrusions a global concern for the entire network; the LIDS existing on different nodes collaborate. Collaboration among the nodes is achieved using two types of data: security data to obtain complementary information from collaborating hosts, and intrusion alerts to inform others of a locally detected intrusion.

A. Mobile IDS Agents

Each node in the network will have an IDS agent running on it all the time. This agent is responsible for detecting intrusions based on local audit data and participating in cooperative algorithms with other IDS agents to decide if the network is being attacked. Each agent has five parts: a local audit trail, a local intrusion database (LID), a secure communication module, anomaly detection modules (ADM), and misuse detection modules (MDM).

B. Local Intrusion Database (LID)

LID is a local database that collects all information necessary for the IDS agent, such as the signature files of known attacks, the established patterns of the users on the network, and the normal traffic flow of the network. The ADM and MDM communicate directly with the LID to determine if an intrusion is taking place.

C. Secure communication module

This is necessary to enable an IDS agent to communicate with other IDS agents on other nodes. It will allow the MDM and ADM to use cooperative algorithms to detect intrusions. It may also be used to initiate a global response when an IDS agent or a group of IDS agents detects an intrusion. Data communicated via the secure communication module needs to be encrypted.

D. Anomaly Detection Modules (ADM)

ADM are responsible for detecting a different type of anomaly. There can be from one to many ADM on each mobile IDS agent, each working separately or cooperatively with other ADM.

E. Anomaly Detection Modules (ADM)

These identify known patterns of attacks that are specified in the LID. Like the ADM, if the audit data available locally is sufficient to determine if an intrusion is taking place, the proper response can be initiated. It is also possible for an MDM to use a cooperative algorithm to identify an intrusion.

F. Misuse Detection Modules (MDM)

These identify known patterns of attacks that are specified in the LID. Like the ADM, if the audit data available locally is sufficient to determine if an intrusion is taking place, the proper response can be initiated. It is also possible for an MDM to use a cooperative algorithm to identify an intrusion.

G. Cooperative Detection

Any node that detects locally a known intrusion or anomaly with strong evidence can determine independently that the network is under attack and can initiate a response. However, if a node detects an anomaly or intrusion with weak evidence, it can initiate a cooperative global intrusion detection procedure. This procedure works by propagating the intrusion detection state information among neighboring nodes.

H. Stationary Secure Database (SSD)

This acts as a secure trusted repository for mobile nodes to obtain information about the latest misuse signatures and find the latest patterns of normal user activity. It is assumed that the attacker will not compromise the SSD, as it is stored in an area of high physical security. The mobile IDS agents will collect and store audit data (user commands, network traffic), while in the field, and will transfer this information when they are attached to the SSD. When the IDS agents are connected to SSD, they will gain access to the latest attack signatures automatically. As this intrusion framework supports reputation mechanism which helps the mobile nodes in optimizing:

- i. Power consumption
- ii. Battery life

I. Power Management

A major challenge to the design of a power management framework for ad hoc networks is that energy conservation usually comes at the cost of degraded performance such as lower throughput or longer delay. A naive solution that only considers power savings at individual nodes may turn out to be detrimental to the operation of the whole network. This framework uses Trust manager, which is an important component in reputation mechanism, decides whether an

incoming requesting node can be trusted and routed. Depending upon on trust value packets may be routed. Thus, unnecessary power consumption is avoided.

J. Bandwidth Utilization

Another important factor that affects the performance of ad hoc node is bandwidth utilization. Malicious nodes constantly requests for forwarding packets. This degrades performance of ad hoc nodes. However, this framework uses path manager, which always forwards the packets for the shortest path and trusted route. Thus bandwidth can be saved.

V. CONCLUSION

This framework uses an intrusion detection, which identifies intrusion at locally and globally. However, by using reputation mechanism, system can optimized the power consumption and bandwidth utilization.

REFERENCES

- [1] Liu Jianxiao 1, Li Lijuan 1 "Research of Distributed Intrusion Detection System Model Based on Mobile Agent" In Proceeding of International Forum on Information Technology and Application, pp.53-57,2009 IEEE
- [2] MO Xiu-liang, WANG Chun-dong , "A Distributed Intrusion Detection System Based on Mobile Agents" In 2nd International conference in Biomedical Engineering and Informatics,2009,IEEE
- [3] Nita Patil ,Chhaya Das, "Analysis of Distributed Intrusion Detection Systems using Mobile Agents"2008 IEEE Jing Xu, Yancheng , Yongzhong Li,"A New Distributed Intrusion Detection Model Based on Immune Mobile Agent" In Asia-Pacific Conference on Information Processing ,2009 IEEE[5] Y. Zhang , W. Lee, "Intrusion Detection in Wireless Ad Hoc Networks," 6th Int'l. Conf. Mobile Comp. And Net. Aug. 2000, pp. 275-83.
- [4] KaKachirski, R. Guha, "Intrusion Detection Using Mobile Agents in Wireless Ad Hoc Networks," Knowledge Media Net., Proc. IEEE Workshop., July 10 , 2002, pp. 153-58.
- [5] R.Raamanujan et al., "Techniques for Intrusion-Resistant Ad Hoc Routing Algorithms(TIARA) " MILCOM 2000, vol. 2, Oct. 22-25, 2000, pp. 660-64.
- [6] Po Wah Yau , Chris J Mitchell," Reputation methods for routing security for mobile ad hoc networks", IEEE 2003, pp-130-137.

Performance Comparison of Assorted Color Spaces for Multilevel Block Truncation Coding based Face Recognition

Dr. H.B. Kekre

Senior Professor

Computer Engineering Department

MPSTME, SVKM's NMIMS

(Deemed-to-be University)

Mumbai, India

Dr. Sudeep Thepade

Associate Professor

Computer Engineering Department

MPSTME, SVKM's NMIMS

(Deemed-to-be University)

Mumbai, India

Karan Dhamejani,

Sanchit Khandelwal,

Adnan Azmi

B.Tech Students

Computer Engineering Department

MPSTME, SVKM's NMIMS

(Deemed-to-be University)

Mumbai, India

Abstract— The paper presents a performance analysis of Multilevel Block Truncation Coding based Face Recognition among widely used color spaces. In [1], Multilevel Block Truncation Coding was applied on the RGB color space up to four levels for face recognition. Better results were obtained when the proposed technique was implemented using Kekre's LUV (K'LUV) color space [25]. This was the motivation to test the proposed technique using assorted color spaces. For experimental analysis, two face databases are used. First one is "Face Database", developed by Dr. Libor Spacek which has 1000 face images and the second one is "Our Own Database" which has 1600 face images. The experimental results showed that Block Truncation Level 4 (BTC-Level 4) gave the best result in every color space. It is observed that the proposed technique functions better in the YIQ color space.

Keywords- Face recognition, Block Truncation Coding, RGB, K'LUV, YIQ, YUV, YCbCr, YCgCb, Multilevel BTC.

I. INTRODUCTION

Face recognition plays an imperative role in identification and for authentication purpose, in our everyday lives. In real time, this identification must be efficient, liable and faster. Face recognition is preferred over other techniques like fingerprint recognition, iris recognition because it does not require explicit cooperation from users. Also special equipments are not required to capture the image [21, 22, 23]. It is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source.

Face recognition can be achieved by comparing the input query face image with the existing face images stored in the database. It is the fastest growing biometric technology. Some of the applications of face recognition include physical, security and computer access controls, law enforcement [12, 13], criminal list verification, surveillance at various places [15], forensic, authentication at airports [17], etc.

A large number of face detection algorithms are derived from algorithmic approach [2, 3, 4, 5, 6, 7, 8, 9, 24] and some image morphological techniques [18]. However most of the works concentrate on single face detection, with some constrained environments. In this paper performance comparison of Multilevel Block Truncation Coding [1] using various color spaces has been carried out on two face databases. Results further revealed that the YIQ color space outperforms all the other color spaces at each stage of Multilevel BTC.

II. BLOCK TRUNCATION CODING AND MULTILEVEL BLOCK TRUNCATION CODING

Block truncation coding (BTC) [11, 12, 13, 14] is a relatively simple image coding technique developed in the early years of digital imaging more than 29 years ago. Block Truncation Coding (BTC) was first developed in 1979 for grayscale image coding [13]. Although it is a simple technique, BTC has played an important role in the history of digital image coding in the sense that many advanced coding techniques have been developed based on BTC or inspired by the success of BTC. It is a straightforward technique which demands very less computational complexity.

In the proposed technique, Multilevel Block Truncation Coding, BTC has been implemented using the RGB color space up till four levels [1, 13]. The feature vector size at BTC-Level 1, BTC-Level 2, BTC-Level 3 and BTC-Level 4 is 6, 12, 24 and 48 respectively. In the same way BTC is implemented on the following color spaces: K'LUV, YUV, YCbCr, YIQ and YCgCb.

III. CONSIDERED COLOR SPACES[12,26,27]

A. Kekre's LUV [25]

K'LUV color space is a special case of Kekre transform. Where L gives luminance and U and V gives chromaticity values of color image. Positive value of U indicates prominence of red component in color image and negative value of V indicates prominence of green component.

Equation (1) gives the RGB to LUV conversion matrix which indicates the corresponding L, U and V components for an image from the R, G and B components.

$$\begin{bmatrix} L \\ U \\ V \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

The reverse conversion, that is from LUV color space to RGB color space is given in (2).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} L/3 \\ U/6 \\ V/2 \end{bmatrix} \quad (2)$$

B. YCbCr

In YCbCr color Space, Y gives luminance and Cb and Cr gives chromaticity values of color image.

Equation (3) gives the RGB to YCbCr conversion matrix which indicates the corresponding Y, Cb and Cr components for an image from the R, G and B components.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.2989 & 0.5866 & 0.1145 \\ -0.1688 & -0.3312 & 0.5000 \\ 0.5000 & -0.4184 & -0.0816 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

The reverse conversion, that is from LUV color space to RGB color space is given in (4).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & -0.0010 & 1.4020 \\ 1 & -0.3441 & -0.7140 \\ 1 & 1.7718 & 0.0010 \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (4)$$

C. YUV

In YUV color space, Y component gives the luminance (brightness) of the color and while U and V components give the chrominance (color).

Equation (5) gives the RGB to YUV conversion matrix which indicates the corresponding Y, U and V components for an image from the R, G and B components.

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.14713 & -0.22472 & 0.436 \\ 0.615 & -0.51498 & 0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

The reverse conversion, that is from YUV color space to RGB color space is given in Equation (6).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0.7492 & -0.50901 & 1.1398 \\ 1.0836 & -0.22472 & -0.5876 \\ 0.97086 & 1.9729 & -0.000015 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \quad (6)$$

D. YIQ

The YIQ color space is derived from YUV colour space. I stands for in phase and Q for Quadrature.

Equation (7) gives the RGB to YIQ conversion matrix which indicates the corresponding Y, I and Q components for an image from the R, G and B components.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.595716 & -0.274453 & -0.321263 \\ 0.211456 & -0.522591 & 0.31135 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (7)$$

The reverse conversion, that is from YIQ color space to RGB color space is given in (8).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.9563 & 0.6210 \\ 1 & -0.2721 & -0.6474 \\ 1 & -1.107 & 1.7046 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \quad (8)$$

E. YCgCb

To get Y, Cg and Cb components we need the conversion of RGB to YCgCb. The RGB to YCgCb conversion matrix is given in (9) gives the Y, Cg, Cb components of color image for respective R, G and B components.

$$\begin{bmatrix} Y \\ Cg \\ Cb \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (9)$$

The YCgCb to RGB conversion matrix given in (10) gives the R, G, B components of color image for respective Y, Cg and Cb components.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = 1/3 \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} Y \\ Cg \\ Cb \end{bmatrix} \quad (10)$$

IV. PROPOSED METHOD

To calculate the feature vector of each image in the database set and the query image, Multilevel Block Truncation Coding has been used for each of the assorted color space.

At each level of BTC, the feature vector of the query image is compared with the feature vector of each image in the training set. The comparison (Similarity measurement) is done by Mean Square Error (MSE) given by equation 11.

$$MSE = \frac{1}{MN} \sum_{y=1}^M \sum_{x=1}^N [I(x, y) - I'(x, y)]^2 \quad (11)$$

Where,

I & I' are two feature vectors of size M*N which are being compared.

False Acceptance Ratio (FAR) and Genuine Acceptance Ratio (GAR) have been used as the performance evaluation parameters to assess the competence of each considered color space.

V. IMPLEMENTATION

A. Platform

The effectuation of the Multilevel BTC is done in MATLAB 2010. It is carried out on a computer using an Intel Core i5-2410M CPU (2.4 GHz).

B. Database

The experiments were performed on two face databases.

1) Face Database [16]

This database is created by Dr Libor consisting of 1000 images (each with 180 pixels by 200 pixels), corresponding to 100 persons in 10 poses each, including both males and females. All the images are captured against a dark or bright homogeneous background, little variation of illumination, different facial expressions and details. The subjects sit at fixed distance from the camera and are asked to speak, whilst a sequence of images is taken. The speech is used to introduce facial expression variation. The images were taken in a single session. The ten poses of Face database are shown in Figure 1.

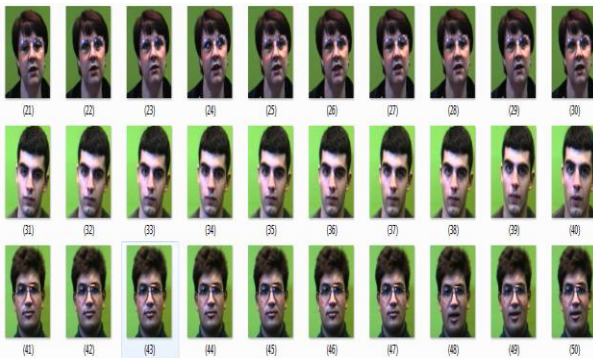


Figure 1: Sample images from Face database

2) Our Own Database [1, 20]

This database consists of 1600 face images of 160 people (92 males and 68 females). For each person 10 images are taken. The images in the database are captured under numerous illumination settings. The images are taken with a homogenous background with the subjects having different expressions. The images are of variable sizes, unlike the Face database. The ten poses of Our Own Database are shown in Figure 2.

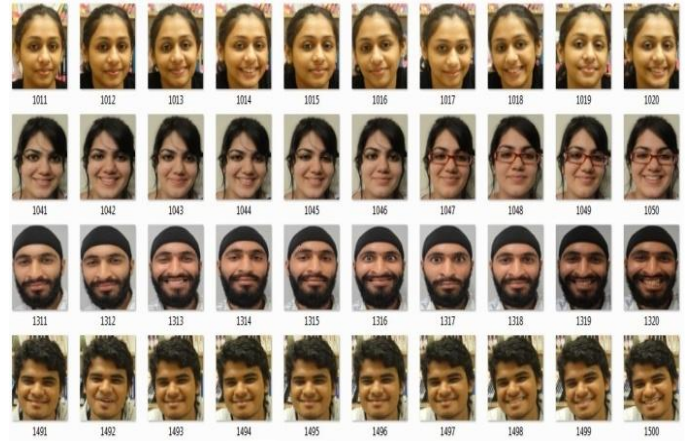


Figure 2: Sample images from Our Own Database

VI. RESULTS AND DISCUSSIONS

False Acceptance Rate (FAR) and Genuine Acceptance Rate (GAR) are standard performance evaluation parameters of face recognition system.

The False acceptance rate (FAR) is the measure of the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user. A system's FAR typically is stated as the ratio of the number of false acceptances divided by the number of identification attempts.

$$FAR = (\text{False Claims Accepted} / \text{Total Claims}) \times 100 \quad (12)$$

The Genuine Acceptance Rate (GAR) is evaluated by subtracting the FAR values from 100.

$$GAR = 100 - FAR \text{ (in percentage)} \quad (13)$$

For each color space, 10000 queries (10 images for each of the 1000 people) are fired on face database and 16000 queries (10 images for each of the 1600 people) are fired on Our Own Database. At the end, average FAR and GAR of all queries in respective face databases are considered for performance ranking of BTC levels and of the color spaces.

For optimal performance the FAR values must be less and accordingly the GAR values must be high for each successive levels of BTC.

A. Face Database

To analyze the performance of proposed algorithm and for performance ranking of color spaces, 10000 queries are fired for each of the assorted color space. For every color space, every BTC level; feature vector of the query image is calculated and compared with the feature vectors of every image in the database. The FAR and GAR values are calculated by employing equations 12 and 13.

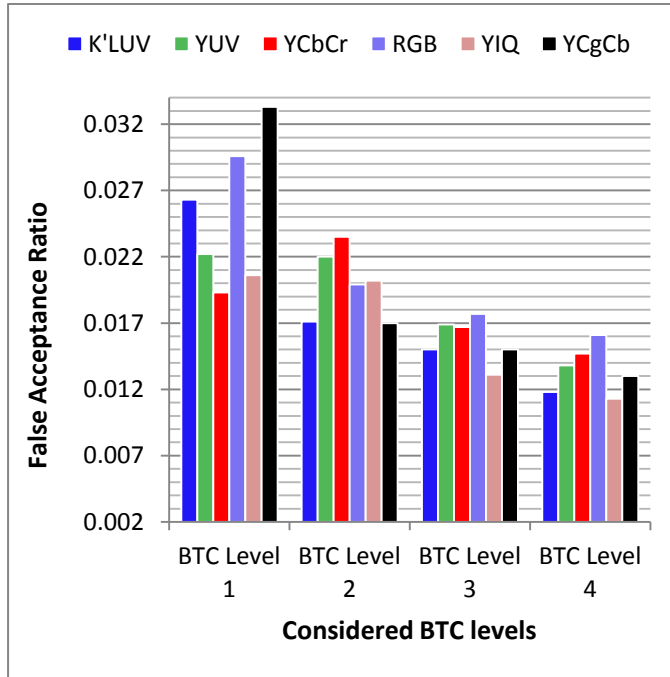


Figure 3. FAR values at different BTC levels of the assorted color spaces for Face Database

Figure 3 gives the FAR values of the different BTC levels based face recognition techniques tested on face database for the considered color spaces. Here it can be seen that the FAR values go on decreasing for each succeeding level of BTC of respective color spaces. This shows that the accuracy of face recognition increases with increasing level of BTC and hence BTC-level 4 gives the best result with the least FAR value in all the color spaces. Also the FAR values of YIQ color space are the least. Thus, it can be concluded that the implementation of BTC levels based face recognition techniques is better when applied in YIQ color space.

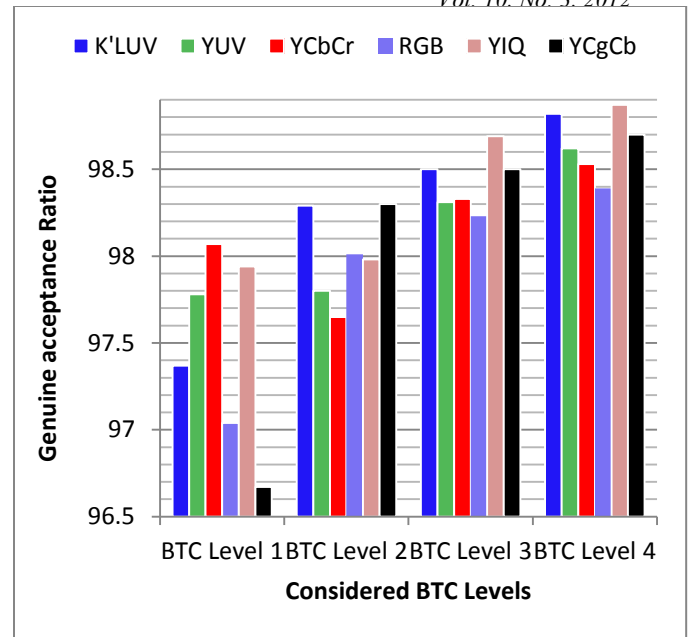


Figure 4. GAR values at different BTC levels of the assorted color spaces for Face Database

Figure 4 gives the GAR values of the different BTC levels based face recognition techniques tested on face database for the assorted color spaces. Here it is observed that with each successive level of BTC the GAR values go on increasing in respective color spaces and hence a BTC-level 4 gives the best result with the highest value in all the color spaces. It is also observed that the YIQ color space shows the highest GAR values at all levels of BTC followed by YCbCr, K'LUV, YUV, YCrgCrb and RGB respectively.

An anomaly is noticed in YCbCr color space for this database. Not conforming to the generally observed pattern, the FAR values increase at the second level of the BTC based face recognition technique.

B. Our Own Database

In all 16000 queries were tested on the database for analyzing the performance of the proposed BTC level based face recognition algorithm for the assorted color spaces. The experimental results of proposed face recognition techniques have shown that BTC level 4 gives the best performance in respective color spaces. The efficiency of the Multilevel BTC based face recognition increases with the increasing levels of BTC.

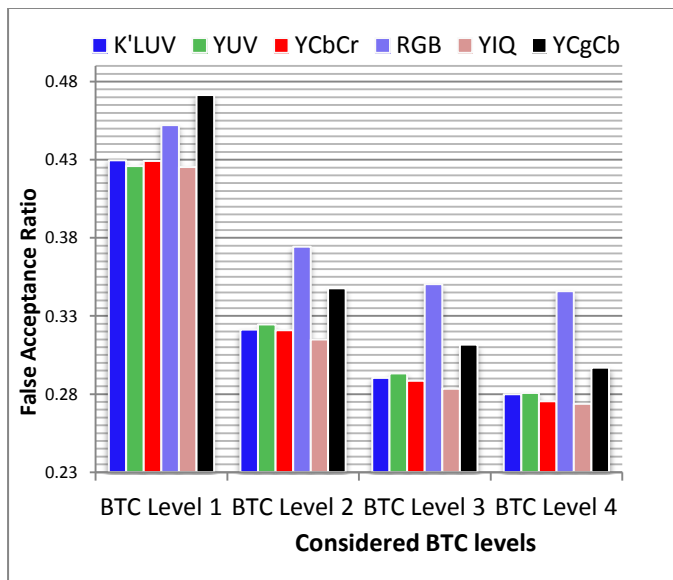


Figure 5. FAR values at different BTC levels of the assorted color spaces for Our Own Database

Figure 5 gives the FAR values of the different BTC levels based face recognition techniques tested on Our Own Database for all color spaces. The FAR values go on decreasing for each succeeding level of BTC of respective color spaces. Thus, when BTC based face recognition techniques is applied on Our Own Database, it gives a result similar to the Face Database; The BTC level 4 gives the best result for respective color spaces and YIQ color space is better than other color spaces for implementing this proposed algorithm.

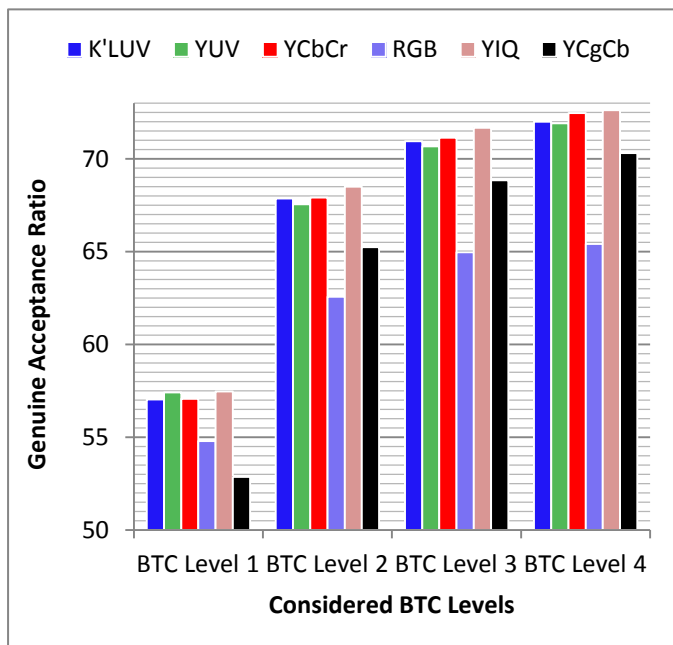


Figure 6. GAR values at different BTC levels of the assorted color spaces for Our Own Database

Figure 6 gives the GAR values of the different BTC levels based face recognition techniques tested on Our Own Database. It can be seen from the above figure that BTC-Level 4 has the highest GAR values and hence it is better than other BTC-Levels. Also the GAR values of YIQ color space are greater than the GAR values of all the other color spaces considered, at all the levels. Thus, it can be concluded that the implementation of BTC levels based face recognition techniques is better when applied in YIQ color space.

VII. CONCLUSION

BTC based face recognition using assorted color spaces have been presented in the paper. Earlier the RGB and K'LUV color spaces were considered and it was observed that better results were shown by the K'LUV color space. In this paper, six color spaces have been considered and the proposed technique has been implemented till four levels of BTC. In all 24 combinations have been tested on two databases; Our Own Database (Not normalized, 1600 face images) and Face Database (Normalized, 1000 face images). It is concluded that the YIQ color space at level four of BTC gives the best results followed by the YCbCr color space at BTC level four.

REFERENCES

- [1] H.B.Kekre, Sudeep D. Thepade, Sanchit Khandelwal, Karan Dhamejani, Adnan Azmi, "Face Recognition using Multilevel Block Truncation Coding" International Journal of Computer Applications (IICA) December 2011 Edition.
- [2] Xiujuan Li, Jie Ma and Shutao Li 2007. A novel faces recognition method based on Principal Component Analysis and Kernel Partial Least. IEEE International Conference on Robotics and Biometrics, 2007. ROBIO 2007
- [3] Shermin J "Illumination invariant face recognition using Discrete Cosine Transform and Principal Component Analysis" 2011 International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT).
- [4] Zhao Lihong , Guo Zikui "Face Recognition Method Based on Adaptively Weighted Block-Two Dimensional Principal Component Analysis"; 2011 Third International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN)
- [5] Gomathi, E, Baskaran, K. "Recognition of Faces Using Improved Principal Component Analysis"; 2010 Second International Conference on Machine Learning and Computing (ICMLC)
- [6] Haitao Zhao, Pong Chi Yuen "Incremental Linear Discriminant Analysis for Face Recognition", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics
- [7] Tae-Kyun Kim; Kittler, J. "Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image" IEEE Transactions on Pattern Analysis and Machine Intelligence, March 2005
- [8] James, E.A.K., Annadurai, S. "Implementation of incremental linear discriminant analysis using singular value decomposition for face recognition". First International Conference on Advanced Computing, 2009. ICAC 2009
- [9] Zhao Lihong, Wang Ye, Teng Hongfeng; "Face recognition based on independent component analysis", 2011 Chinese Control and Decision Conference (CCDC)

- [10] Yunxia Li, Changyuan Fan; "Face Recognition by Non negative Independent Component Analysis" Fifth International Conference on Natural Computation, 2009. ICNC'09'.
- [11] Yanchuan Huang, Mingchu Li, Chuang Lin and Linlin Tian. "Gabor-Based Kernel Independent Component Analysis on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP).
- [12] H.B.Kekre, Sudeep D. Thepade, Varun Lodha, Pooja Luthra, Ajay Joseph, Chitrangada Nemani "Augmentation of Block Truncation Coding based Image Retrieval by using Even and Odd Images with Sundry Color Spaces", Int. Journal on Computer Science and Engg. Vol. 02, No. 08, 2010, 2535-2544
- [13] H.B.Kekre, Sudeep D. Thepade, Shrikant P. Sanas, "Improved CBIR using Multileveled Block Truncation Coding", International Journal on Computer Science and Engineering Vol. 02, No. 08, 2010, 2535-2544
- [14] H.B.Kekre, Sudeep D. Thepade, "Boosting Block Truncation Coding using Kekre's LUV Color Space for Image Retrieval", WASET International Journal of Electrical, Computer and System Engineering (IJCSE), Volume 2, Number 3, pp. 172-180, Summer 2008.
- [15] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Augmented Block Truncation Coding Techniques", ACM International Conference on Advances in Computing, Communication and Control (ICAC3-2009), pp. 384-390, 23-24 Jan 2009, Fr. Conceicao Rodrigues College of Engg., Mumbai
- [16] Developed by Dr. Libor Spacek. Available Online at: <http://cswww.essex.ac.uk/mv/otherprojects.html>
- [17] Mark D. Fairchild, "Color Appearance Models", 2nd Edition, Wiley-IS&T, Chichester, UK, 2005. ISBN 0-470-01216-1
- [18] Rafael C. Gonzalez and Richard Eugene Woods "Digital Image Processing", 3rd edition, Prentice Hall, Upper Saddle River, NJ, 2008. ISBN 0-13-168728-X. pp. 407-413.S
- [19] Dr.H.B.Kekre, Sudeep D. Thepade and Shrikant P. Sanas, "Improved CBIR using Multileveled Block Truncation Coding", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 07, 2010, 2471-2476
- [20] Dr. H.B.Kekre , Sudeep D. Thepade and Akshay Maloo, "Face Recognition using Texture Features Extracted from Walshlet Pyramid ", Int. J. on Recent Trends in Engineering & Technology, Vol. 05, No. 01, Mar 2011.
- [21] Koji kotani, Chen Qiu and Tadahiro Ohmi, "Face Recognition Using Vector Quantization Histogram Method". International Conference on Image Processing, Volume 2, pp.105-108,2002.
- [22] Shang-Hung Lin, "An Introduction to Face Recognition Technology", Informing Science Special Issue on Multimedia Informing Technologies- Part 2, Volume 3 No.1, 2000.
- [23] H. B. Kekre, Sudeep Thepade, Akshay Maloo, "Eigenvectors of Covariance Matrix using Row Mean and Column Mean Sequences for Face Recognition", CSC-International Journal of Biometrics and Bioinformatics (IJBB), Volume (4): Issue (2), pp. 42-50, May 2010.
- [24] H. C. Vijaya Lakshmi, D. Patil Kulakarni "Segmentation algorithm for multiple face detection in color images with skin tone regions using color spaces and edge detection techniques," International journal of computer theory and engineering 1793-8201,2010.
- [25] Dr. H. B. Kekre, Sudeep Thepade, Karan Dhamejani, Adnan Azmi, Sanchit Khandelwal, "Improved Face Recognition with Multilevel BTC using Kekre's LUV Color Space", IJACSA
- [26] Dr. H. B. Kekre, Sudeep Thepade, Adib Parkar "A Comparison of Haar Wavelets and Kekre's Wavelets for Storing Colour Information in a Greyscale Image" International Journal of Computer Applications (0975 - 8887) Volume 11- No.11, December 2011.
- [27] Dr. H. B. Kekre, Sudeep Thepade, Nikita Bhandari "Colorization of Greyscale Images Using Kekre's Biorthogonal Color Spaces and Kekre's Fast Codebook Generation Advances in Multimedia" An International Journal (AMIJ), Volume (1): Issue (3)

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engineering. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. Now he is Senior Professor at MPSTME, SVKM's NMIMS University. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E./B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 350 papers in National / International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE Recently ten students working under his guidance have received best paper awards and two have been conferred Ph.D. degree of SVKM's NMIMS University. Currently 10 research scholars are pursuing Ph.D. program under his guidance.

Dr. Sudeep D. Thepade has Received B.E.(Computer) degree from North Maharashtra University with Distinction in 2003, M.E. in Computer Engineering from University of Mumbai in 2008 with Distinction, Ph.D. from SVKM's NMIMS (Deemed to be University) in July 2011, Mumbai. He has more than 08 years of experience in teaching and industry. He was Lecturer in Dept. of Information Technology at Thadomal Shahani Engineering College, Bandra(W), Mumbai for nearly 04 years. Currently working as Associate Professor in Computer Engineering at Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS (Deemed to be University), Vile Parle (W), Mumbai, INDIA. He is member of International Advisory Committee for many International Conferences, acting as reviewer for many referred international journals/transactions including IEEE and IET. His areas of interest are Image Processing and Biometric Identification. He has guided five M.Tech. projects and several B.Tech projects. He has more than 130 papers in National/International Conferences/Journals to his credit with a Best Paper Award at International Conference SSPCCIN-2008, Second Best Paper Award at ThinkQuest-2009, Second Best Research Project Award at Manshodhan 2010, Best Paper Award for paper published in June 2011 issue of International Journal IJCSIS (USA), Editor's Choice Awards for papers published in International Journal IJCA (USA) in 2010 and 2011.

Sanchit Khandelwal is currently pursuing B.Tech. (CE) from MPSTME, NMIMS University, Mumbai. His areas of interest are Image Processing and Computer Networks and security. He has 2 paper in an international journal to his credit.

Karan Dhamejani is currently pursuing B.Tech. (CE) from MPSTME, NMIMS University, Mumbai. His areas of interest are Image Processing, Computer Networks and UNIX programming. He has 3 papers in an international journal to his credit.

Adnan Azmi is currently pursuing B.Tech. (CE) from MPSTME, NMIMS University, Mumbai. His areas of interest are Image Processing and Computer Networks. He has 2 paper in an international journal to his credit.

Step Tapered waveguide with cylindrical waveguide

Harshukumar Khare,
M.E (EXTC) Final year student
TEC, Nerul, Navi-Mumbai
harshukhare@gmail.com

Prof. R.D. Patane
Asst. Proffessor (EXTC)
TEC, Nerul, Navi-Mumbai
rpatne@yahoo.co.in

Abstract: Tapered Waveguide is a waveguide in which a physical or electrical characteristic changes continuously with distance along the axis of the waveguide. Tapered waveguide offer an excellent means of converting microwave mode sizes to connect Microwave devices of different cross-sectional dimensions. This paper discusses the waveguide component for interconnecting rectangular and circular waveguide using step tapering. Model is designed for the frequency range from 2 to 4 GHz. Dominant Mode conversions ie from TE₁₀ to TM₁₁ is considered for tapering techniques. Step tapering is studied at different step sizes 4mm to 10 mm and analysis is done. All simulations done with CST Microwave studio and S parameters and E field parameters are analyzed. Simulation result shows that wave is properly propagated with no power reflection and low power loss.

Key words: Single & Double Step Tapering, Cylindrical waveguide, CST, S parameter, E Field

I. INTRODUCTION

A rectangular waveguide supports TM and TE modes but not TEM waves. A rectangular waveguide cannot propagate below some certain frequency. This frequency is called the cut-off frequency.

Circular waveguides offer implementation advantages over rectangular waveguide in that installation is much simpler when forming runs for turns and offsets - particularly when large radii are involved and the wind loading is less on a round cross-section, meaning towers do not need to be as robust. Manufacturing is generally simpler, too, since only one dimension the radius needs to be maintained. Applications where differential rotation is required, like a rotary joint for a radar antenna, absolutely require a circular cross-section, so even if rectangular waveguide is used for the primary routing, a transition to circular and then possibly back to rectangular is needed. Calculations for circular waveguide require the application of Bessel functions, so working equations with a cheap calculator is not going to happen. However, even spreadsheets have Bessel function capability nowadays, so determining cutoff frequencies, field strengths, and any of

the other standard values associated with circular waveguide can be done relatively easily.

A waveguide taper can always be built to have as low a mode conversion as is wanted in a certain frequency band merely by making it long enough. However, an optimally designed taper has the smallest possible length for a given difference in diameters at its two ends for a specified unwanted mode level in a given frequency band. Tapered waveguide for matching impedance is nothing but a tapered waveguide in which only one mode is propagating. Power can only be converted into reflected waves, and it is this reflected power which is kept small in a properly designed transmission line taper. If more than one mode is propagating, power will be scattered not only into the reflected wave but also into the other propagating modes. In fact, the power scattered into backward traveling waves is quite small compared to the power scattered into forward traveling waves, and only the latter need be considered in a multimode waveguide taper. Therefore, the mode conversion in the waveguide transition corresponds to the reflection in transmission line taper.

A waveguide mode is a unique arrangement of the electric and magnetic fields propagating in the z-direction that satisfies all Maxwell equations and boundary conditions imposed by the geometry of the conductors of the transmission system. Various waveguide modes are TEM, TE, TM and Hybrid modes. Dominant mode in Rectangular waveguide is TE₁₀ and in circular waveguide TE₁₁. To convert dominant mode in rectangular waveguide to dominant mode in circular waveguide tapered waveguide is used. There are different types of tapering such as step tapering, conical tapering elliptical tapering, etc. Analysis has been done using Step tapering with CST Microwave Studio.

II. DESIGN ASPECT

The simulation was done by Transient solver of CST Microwave Studio. The Cartesian coordinate system (x, y, and z) is used to model the 3D structure. Design & analysis has been done with tapering and without tapering.

A. Design without tapering

The rectangular brick is directly connected to circular waveguide. Port of rectangular brick (Port 2) is excited and the S parameters are obtained by Transient Solver. 3D model of cylindrical waveguide and Rectangular Brick Without tapering is shown in fig 1.

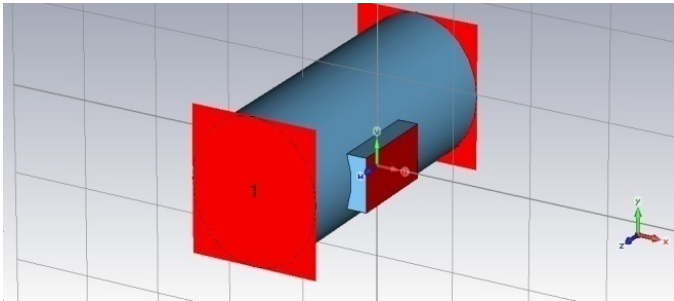


Fig 1 - 3D model of cylindrical waveguide and Rectangular Brick Without tapering

S11 plot gives that around 2.1 GHz S11 goes upto 19 dBs shown in Fig 2 which is not a desirable case as huge amount of power is reflected back to the source damaging the network analyzer.

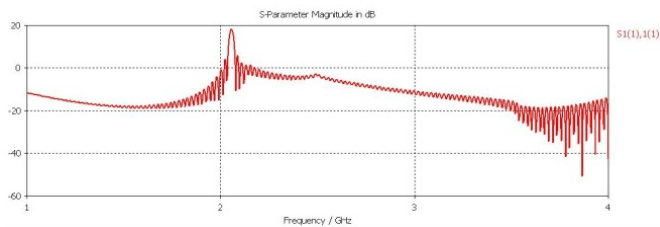


Fig 2 - S11 plot

S21 plot is shown in fig. 3 which indicates that no sufficient output is coupled to the output port as it is approaching to 0dB i.e. no power is coupled from port 1 to port 2 and vice versa.

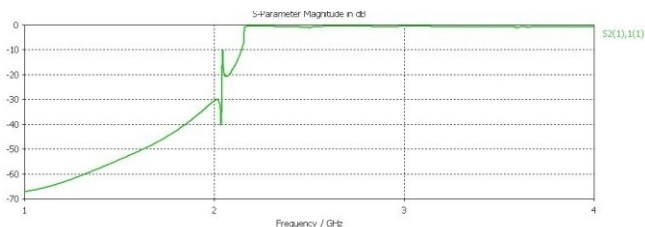


Fig 3 - S21 plot

B. Design with tapering

i) Single Step tapering

The 3D model of single step tapering is shown in Fig 4.

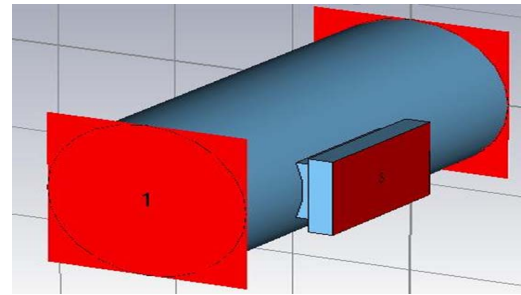


Fig 4 - 3D model of cylindrical waveguide and Rectangular Brick with Step tapering

Simulated results for S parameters & E field are calculated for different step sizes for sizes of 4 mm & 10mm their S11 and S21 are shown in Fig 5.

It is seen from Fig 5 that for step size 4 mm plots the sufficient amount of power is reflected back which shows that taper is not properly coupled and for step size 10 mm it is seen that no power is reflected back hence source is safe in this case.

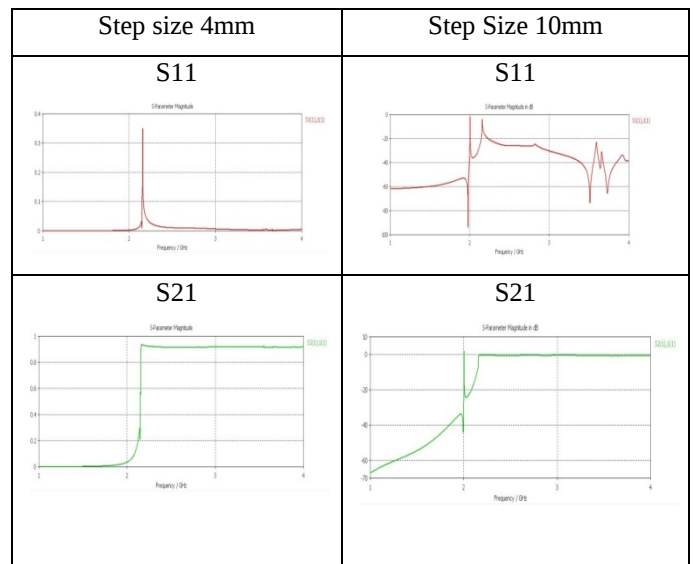


Fig-5- S11 & S21 for Step size 4mm & 10 mm

E Field distribution in single step tapered and cylindrical waveguide is shown in Fig. 6.

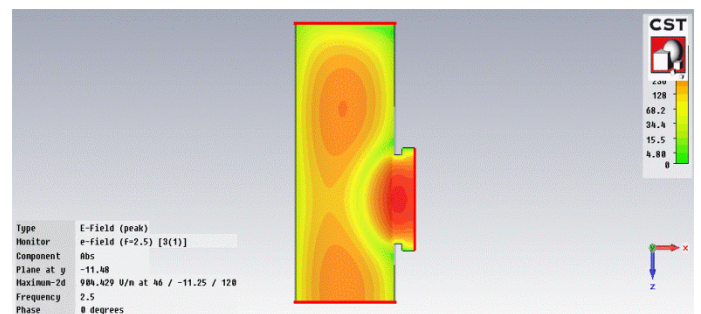


Fig-6 E Field distribution in single step tapering

ii) Double Step Tapering

The schematic of double step tapering is shown in Fig 7 in which two steps of 10 mm are used.

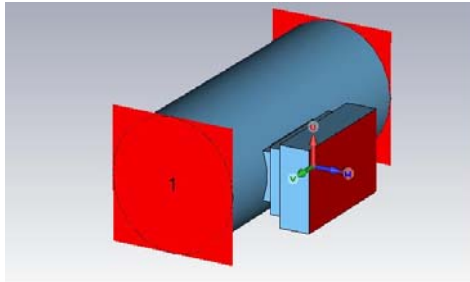


Fig 7 - Double Step Tapering

Their S11, S21 plots and E field distribution are shown in Figs 8, 9 & 10. In double step tapering from Fig 8 & 9 it shows that impedance matching problem occurs. Due to this wave is not guided properly from cylindrical waveguide to rectangular waveguide.

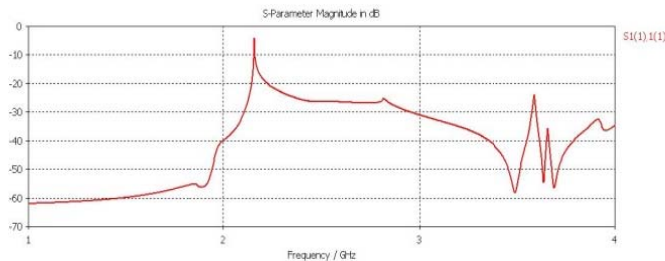


Fig 8 - S11 plot for double step tapering

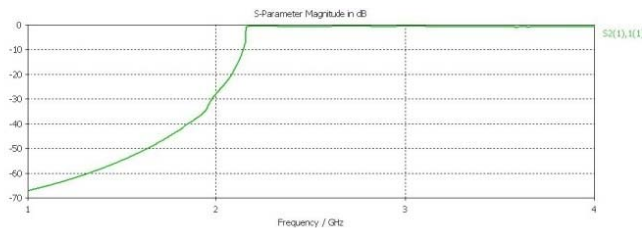


Fig 9 - S21 plot for double step tapering

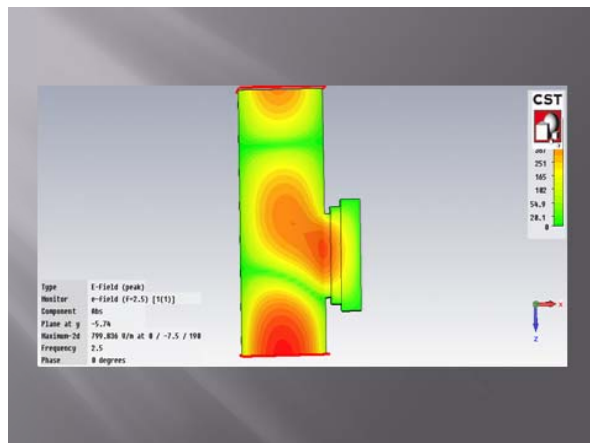


Fig 10 - E Field Distribution in Double step tapering

III. RESULT & CONCLUSION

The total simulation process was done by CST Microwave Studio. It is infer that Tapering between two waveguides is good possible solution to connect two waveguides. Tapered waveguide offer an excellent means of converting microwave mode sizes to connect microwave devices of different cross-sectional dimensions. Properly wave is guided in single stepping with 10 mm step size with no reflection in waveguide. Impedance matching is the minor problem with double step.

ACKNOWLEDGMENT

The authors are grateful to Dr. R.C Sethi, HOD,EXTC TEC and Prof. Mrs. Jyothi Digge, PG coordinator, EXTC,TEC for their great support and valuable guidance. They would also like to acknowledged help & support received from Dr. Abhay Deshpande, Scientist, SAMEER,IIT-B.

REFERNCES

- [1] Chen Huaibi, Huang Yuanzhong, Lin Yuzheng, Tong Dechun, Ding Xiaodong Department of Engineering physics, Tsinghua University, Beijing 100084, BACKWARD TRAVELING WAVE ELECTRON LINAC, 1998 IEEE
- [2] J. Petillo, W. Krueger, A. Mondelli, "Frequency Domain Determination of the Waveguide Loaded Q for the SSCL Drift Tube Linac" IEEE Particle accelerator conference 1993.
- [3] Muralidhar Yeddulla , Sami Tantawi , SLAC, Menlo Park, "Analysis of a Compact Circular TE_{0,1} - Rectangular TE_{0,2} Waveguide Mode Converter", Proceedings of PAC07, Albuquerque, New Mexico, USA, 2007,pp-587-589
- [4] L. Solymar, "Spurious mode generation in nonuniform waveguide," IRE Transactions on Microwave Theory and Techniques, vol. MTT-7, pp. 379-383, 1959.
- [5] Dr. R.C.Sethi etc," Design of RF structure for 10 MeV,10 KW, Industrial RF electron linac.
- [6] P. K. Jana, Purushottam Shrivastava, Nita. S. Kulkarni, "Design of Microwave Coupler for 10 MeV Electrons LINAC"

A Panoramic Approach on Software Quality Assurance Proposed By CMM and XP

CH.V. Phani Krishna^{*1}
Associate professor,
CSE Department,
KL University, Guntur dt., India.

Dr.G.Rama Krishna^{*2}
Professor,
CSE Department,
KL University, Guntur dt., India.

Dr. K.Rajasekhara Rao^{*3}
Dean of student and faculty welfare,
KL University, Guntur Dt., India.

Abstract---The main objective of this paper is to compare Capability Maturity Model (CMM) and Extreme Programming (XP) regarding their software quality support in terms of software quality development. The main goal is to analyze or measure how the code is framed for particular software, and apply software to show the result.

KEYWORDS—Sqa, Xp, Cmm.

I. INTRODUCTION

The software quality engineering focuses on the processes involved in the development and establishment of software quality. Software quality engineering includes software quality development and software quality assurance. Software quality development consists of requirements engineering, system and software design and implementation. Software quality assurance consists of software quality assurance, quality management and verification and validation. Software quality is achieved by three approaches: testing and static analysis and development approaches. The integration of all three approaches is the most desirable approach.

Different users think differently about the quality of software. The end-user expects the software to help him to do the job faster and easier with adequate help. The buyer expects the software to meet the specifications within the contract terms. The developer attempts to trace defects and focuses faster development as well as higher productivity. The maintainer expects software to be understandable, testable, and modifiable, with all documentation.

The characteristics of software quality in product transition are reusability, portability and interoperability. The characteristics of software quality in product revision are maintainability, adaptability and expandability. The characteristics of software quality in product operation are usability, security, efficiency, correctness and reliability. The attributes of software quality are manageability, efficiency, safety, expandability, reliability, flexibility and usability.

There are quantitative as well as qualitative benefits in maintaining quality assurance. The Quantitative benefits are reduced costs, greater efficiency, better performance, less unplanned work and fewer disputes. The Qualitative benefits are improved visibility and predictability, better control over contracted products, improved customer

confidence, better quality, problems show up earlier and reduced risk.

II. SOFTWARE QUALITY ASSURANCE PROPOSED BY CMM:

It is well known the CMM describes an evolutionary improvement path to a mature disciplined process.

CMM defines key practices to improve the ability of the organization to meet goals for cost, functionality and quality. SQA activities are defined at level 2

According to CMM the purpose of software quality assurance (SQA) is to provide the management with appropriate visibility into the process being used by the software project and of the products being built. It is required that the project follows a return organizational policy for implementing the SQA.

CMM defines eight activities to be performed as follows:

- ❖ A SQA plan is prepared for the software project according to documented procedure.
- ❖ SQA's group activities includes:
- ❖ Responsibilities and authority of SQA group
- ❖ Resource requirements of SQA group
- ❖ Schedule and funding of the project.
- ❖ Participation in establishing the software development plan (SDD).
- ❖ Evaluations to be performed.
- ❖ Audits and reviews to be conducted.
- ❖ Projects standards and procedures forming basis for SQA reviews.
- ❖ Procedures for documenting and tracking non-Compliance issues.
- ❖ Documentation to produce.
- ❖ Method and frequency to provide feedback to other related group.

- ❖ The SQA group participates in the preparation and review of the project's software development plan, standards and procedures and audit the software project.

- ❖ The SQA group audits designated software work products to verify compliance.

- ❖ The SQA group periodically reports the result of its activities to the software engineering group.

- ❖ Deviations identified in the software activities and software work products are documented and handled according to documented procedure.

- ❖ The SQA group conducts periodic reviews of its activity and findings with customers SQA personnel as appropriate.

III. CMM LEVELS KEY PROCESS AREAS AND THEIR PURPOSE:

A. Initial:

This is the starting point for use of a new or undocumented, repeated process. Little documentation is necessary if any processes and procedures take place. Success is only achieved by the heroic actions of team members.

When to use:

Used for a kind projects of very limited scope.

B. Repeatable:

The process is at least documented sufficiently such that repeating the same steps may be exempted. Enough documentation exists that the QA process is repeatable.

When to use:

This is used for any project that will be done again, whether as an upgrade or a somewhat similar variation.

C. Defined:

The process is defined/confirmed as a standard business process, and decomposed to levels 0, 1 and 2 (the latter being Work Instructions). QA documentation and processes & procedures are standardized. Templates exist for all documentation and a QA "system" exists.

When to use:

This is critical for a QA department that must provide QA for multiple projects. This avoids reinventing the wheel for each project.

D. Managed:

The process is quantitatively managed in accordance with agreed-upon metrics. The exact time & resources required to provide adequate QA for each product is known precisely so that timetables and quality levels are met consistently.

When to use:

This requires an existing data set based on previous QA projects. This level can only be achieved by well documented experience.

E. Optimizing:

Process management includes deliberate process optimization / improvement. QA processes and procedures are understood well enough to be refined and streamlined.

When to use:

This should be actually used in every stage. In Level 5, this is the only thing left to work on.

It would be enlightening to conduct a CMM assessment of a team successfully practicing XP. In fact, XP team would achieve a maturity level 2 or better. CMM level 2 is about managing project requirements and schedules effectively and repeatedly. XP claims to do just that, using story cards and a planning game [4].

Thus, the software engineering goals are worthy and they can even be implemented with lightweight methodologies where appropriate. XP is compatible to CMM as well. Software quality assurance consists of Software quality assurance, quality management and verification and validation [5]. Software quality is achieved by three approaches: Testing, Static analysis and development approach. The integration of all the three approaches is the most desirable approach. A different categorization of approaches towards software quality regards four ways to establish software quality: Software quality via better quality evaluation, better measurement, better processes and better tools [6].

Large-scale quality models like Capability Maturity Model (CMM) or ISO-9001 tend to form a SQA in terms of a "process police". [7] SQA takes care only that the process requirements are met but does not consider the quality of the process itself. Instead of SQA in terms of CMM or ISO 9001 a better solution is to embed quality evaluation in the development process.

XP require certain adaptations in order to fulfill CMM requirements specialized maturity models for XP are introduced by combining Capability Maturity Model (CMM) with Personal Software Process (PSP) [8, 3]. Therefore, instead of eliciting SQA in terms of CMM a better solution can be embedded for quality evaluation in XP [9, 10].

IV. SOFTWARE QUALITY ASSURANCE PROPOSED BY XP:

A. Iterative Software Development:

To establish higher software quality, a software development process has to use an iterative and incremental development approach. By using iterative approach a process can gain more flexibility in dealing with changing requirements or scope. The Short Releases of the product force early feedback from the customer as well as stakeholders which is important for improvement of overall

quality of the software. XP builds on a very strict iterative approach limiting the time needed to encounter errors and forces developers to fix the problem as soon as possible.

B. *Quality as a Primary Objective:*

XP software development process defines quality as a major objective to improve the overall quality of the software. Quality targets have to be defined by involving project team members and customer (On-Site Customer). Thus the quality goals become achievable and measurable.

C. *Continuous Verification of Quality:*

This includes extensive testing. Besides internal unit testing, external acceptance tests with the customer are needed too, in order to verify that the product fulfills the needs and requirements of the customer (Test-Driven Development).

D. *Customer Requirements:*

The requirements of the customer who normally does not have a deep technical knowledge have to be considered, so that developers are able to build an application based on that information. Thus it is necessary that the project team understands the customer and his business. Otherwise it is not possible to implement the customer needs accurately. XP teams focus on the customer needs and requirements throughout the entire project by means of communication and by framing user stories.

E. *Architecture Driven:*

Architecture of a system has a major impact on the overall quality of the product. Using a simple well-designed architecture allows easy integration and reuse (Simple Design and Continuous Integration).

F. *Focus on Teams:*

Focusing on team work also effects the motivation of project members. Seeing everyone as an equally important part of the project leads to a high identification of the team members with the product. Hence the project code is not owned by any single programmer but owned by the team collectively (Collective Code Ownership).

G. *Pair Programming:*

Better solutions are more likely with Pair Programming since two persons most likely have different perspectives of the same problem and therefore they complement each other in solving it. This approach saves time and minimizes the number of errors. This is an explicit practice of XP.

H. *Tailoring with Restrictions:*

Software development process should rely on core elements. Building on these core elements the process should adapt practices (tailoring) according to the project type and project size (eg. RDP)

I. *Risk management:*

Risk management enables early risk mitigation and the possibility to act instead of to react to problems and risks. A well-defined risk awareness and mitigation management form together an effective risk management and is a key factor in achieving high product quality.

V. EXISTING SYSTEM

In the existing system, a large number of codes are divided into only two modules. So in the existing system, performance analysis takes more time and is also not accurate.

As per Mancoridis et al., the earliest of software metrics deal with the measurement of code complexity and its maintainability. He measured the Modularization Quality (MQ) which is the combination of coupling and cohesion. Cohesion is measured as the ratio of the number of internal function-call dependencies that actually exist, to the maximum possible internal dependencies. Coupling is measured as the ratio of the number of actual external function-call dependencies between the two subsystems, to the maximum possible number of such external dependencies. The system level MQ is calculated as the difference between the average cohesion and the average coupling.

VI. PROCESS OF PROPOSED SYSTEM:

In the proposed system, we have considered the leaf nodes of the directory hierarchy of the original source code to be the most fine-grained functional modules. All the files (and functions within) inside a leaf level directory are considered to belong to a single module, with the module corresponding to the directory itself. In this manner, all leaf level directories form the module set for the software.

- A lot of work has been done in the past on automatic approaches for code reorganization. There are certain principles, which are most applicable to code reorganization. Our current ongoing effort is targeted on the reorganization of legacy software, containing millions of lines of non-object oriented code. This code was never modularized, or the modularization was very poor. The problem could be attributed as reorganization of millions of lines of code into modules. This code could reside in thousands of files, in hundreds of directories. Here, each module is formed by grouping a set of entities like files, functions, data structures and variables into a logically interconnected unit.

- Modularization is based on certain design principles:

Principle1: Principles Related to Similarity of Purpose

A module is a cluster of a set of data structures and functions that together offer a distinct purpose. To rephrase, the structures used for representing knowledge and any associated functions in the same module should fit together on the basis of similarity-of-service as opposed to, for instance, on the basis of function call dependencies. Clearly,

every service is related to a specific purpose. The following principles are presented as coming under the “Similarity of Purpose” rubric:

- Maximization of Module Coherence on the Basis of Similarity and Singularity of Purpose
- Minimization of Purpose Dispersion
- Maximization of Module Coherence on the Basis of Commonality of Goals
- Minimization of Goal Dispersion.

Principle 2: Principle Related to Module Compilability

• A universal basis of inter module compilation dependency is that a file from one module needs, through *import* or *include* declarations, one or more files from a different module. As software systems evolve and some modules seem like utilities to developers, it is very easy for such interdependencies to become circular. For apparent reasons, these compilation inter-dependencies make it difficult for modules to grow in parallel, and be tested independently. Hence, as far as possible, it must be possible to compile each module independently of the other modules.

Principle 3: Principle Related to Module Extendibility

One of the most important reasons for object-oriented software development is that the classes can be easily extended whenever one wants a more specialized functionality. Extending object-oriented software through the idea of sub-class allows for a more ordered approach to software development and maintenance, since it makes code authorship and its responsibility easy to identify. While module-level compartmentalization of code does not follow the types of software extension rules that are easy to implement in object-oriented approaches, one nevertheless wants the modules to have similar properties when it comes to code extension and enhancement. The following principle takes into account these aspects of code modularization:

- Maximization of the Stand-Alone Module
- Extendibility

Principle 4: Principle Related to Module Testability

Testing is a vital part of software development. At the most, testing must make sure that software conforms to the existing standards and protocols. This kind of testing is mostly called requirements-based testing. But, most important, testing must guarantee that the software code must act as expected for a whole variety of inputs, both correct and incorrect, and at multiple levels. These levels constitute the level of program at the individual function, and at module interactions level. Testing must account for variety of competencies of all causes that interact with the software. Testing procedures can encounter combinatorial problems if the modules cannot be tested independently. This means that if each module is tested for X inputs, then two inter-dependent modules need to be tested for X^2 inputs.

A modularization procedure must adhere to accomplish the following principle:

- Maximization of the Stand-Alone Testability of Modules

Principle 5: Principles Related to Module Size

When a new software development is started afresh, one cannot have all the modules to be of the same size, and equal to some pre-decided number. Nevertheless, when the modularizing legacy code is completely unorganized, it is essential to be able to bias a clustering algorithm to produce modules of approximately the same size, and whose value depend on considerations which are related to software maintenance.

Putting the whole code in a single module is theoretically a correct modularization, though not a useful one. Hence, we need metrics that can maneuver a modularization algorithm away from making very large modules, towards making modules in the same size, while at the same time also ensure that other considerations are not violated. The following two principles deal with this necessity:

- Principle of Observance of Module Size Bounds
- Principle of Maximization of Module Size

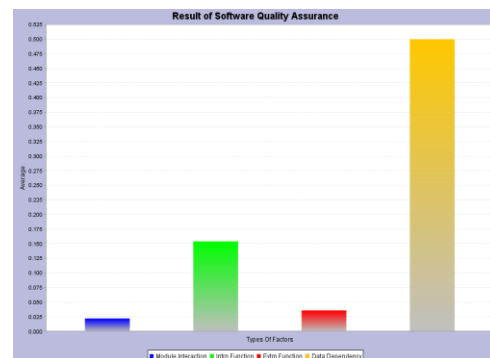


FIG1. Result of Software Quality Assurance by CMM

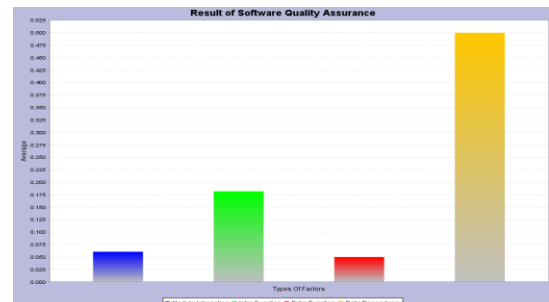


FIG2. Result of Software Quality Assurance by XP

CONCLUSION:

Thus, Practices of XP support software quality development as well as software quality assurance. XP require certain adaptations in order to fulfill CMM requirements specialized maturity models for XP are introduced by combining Capability Maturity Model (CMM) with Personal Software Process. However, much software quality support is implicitly present in XP principles.

REFERENCES:

- [1] B.W.Boehm. **Software Engineering Economics**. Prentice Hall, Englewood Cliffs, NJ, 1981.
- [2] Ward, W.A., and Venkataraman.B, **Some observations on Software quality**, in proceedings of the 37th annual southeast regional conference (CD-ROM), ACM, 1999, Article No.2.
- [3] Microsoft Cooperation: **Microsoft Solutions Framework White Paper**, Microsoft Press, 1999.
- [4] Huo, M., Verner, J., Zhu, L., Babar, M.A: **Software quality and agile methods**. In proceedings of COMPSAC 04, IEEE Computer Soc., 2004, pp.520-25.
- [5] Paulk, N.C: **Extreme Programming from a CMM Perspective**. IEEE software, vol. 18, no.6, IEEE, Nov-Dec.2001, pp.19-26.
- [6] Nawrocki,J.,Walter, B.,and Wojciechowski, A.: **Toward maturity model for Extreme Programming**: In proceedings Euromicro Conference, 2001.IEEE,2001,pp. 233-9.
- [7] Baker, E.B., **Which way, SQA?** .IEEE-Software, vol.18, no.1; Jan.-Feb. 2001; pp. 16-18.
- [8] ManZoni, L.V.; Price, R.T.: **identifying extensions required by RUP(Rational Unified Process) to comply with CMM (Capability Maturity Model) level 2 and 3**. IEEE Transaction on Software Engineering, Vol 29, no.2, IEEE, Feb.2003,pp.181-192.
- [9] Pollice, G.: **Using Rational Unified Process for small Projects: Expanding Upon Extreme Programming**. A Rational Software White Paper, Rational, 2001.
- [10] Runeson, P., Isacsson, P.:**Software Quality Assurance Concepts and Misconceptions**, In Proceedings of the 24th EUROMICRO Conference, IEEE Computer Soc, 1998, pp.853-9.
- [11] Osterweil, L.J.: **Improving the quality of software quality determination processes**, In the Proceedings of the IFIP TC2/WG2.5 Working Conference on Quality of Numerical Software. Assessment and Enhancement, Chapman & Hall, London, 1997, pp.90-105.

AUTHORS PROFILE



Ch.V.Phani Krishna is an Associate Professor in Computer Science and Engineering at KL University. Having more than 10 years of teaching and research experience, he is actively engaged in the research related to Software Engineering. He published 14 International journals. Having Life Membership of ISTE, CSI, IACSIT.



Dr. K.RAJASHEKARA RAO is a Professor of Computer Science and Engineering at KL University and presently holding several key positions in KL University, as Dean (Faculty & Student Affairs) & Principal, KL College of Engineering (Autonomous).Having more than 25years of teaching and research experience, Prof. Rao is actively engaged in the research related to Embedded Systems, Software Engineering and Knowledge Management. He had obtained Ph.D in Computer Science & Engineering from Acharya Nagarjuna University (ANU), Guntur, Andhra Pradesh and produced 35 publications in the International/National Journals and Conferences.

He has been adjudged as best teacher and has been honored with “Best Teacher Award”, six times. Dr. Rao is a Fellow of IETE, Life Member’s of IE, ISTE, ISCA & CSI (Computer Society of India). He has been the past Chairman of the Koneru Chapter of CSI. Presently, Prof. K.R.Rao is the CSI State Student Coordinator of Andhra Pradesh.

Developing Multi-platform package for Remote System Administration

Rawaa Putros Polos Qasha
Department of Computers Sciences
College of Computer Sciences and Mathematics
University of Mosul
Mosul, Iraq
rawa_qasha@yahoo.com

Abstract – This paper presents Multi-Platform System Administration (MPSA) software for administrating and controlling different operating systems such as Linux and Windows system, based on XML-RPC technique and Python libraries.

MPSA software consists of two distinct components: Administration server and administration agent. The server components, work on Linux system, are responsible for communicating with the agents, sending the queries, and retrieve the performance and status data from them. Administration agent, that can be working either in Linux or Windows system, is going to proxy the server requests to build functions and pass the results back to server.

MPSA introduces many services, such as processes management, resource management, gathering system information, system booting, and file browser, by using the advantages of XML-RPC.

MPSA services were implemented to work on the background at the administration Agents, so that the software works in a hidden manner without needing to agents permission or attention.

Keywords- Remote System Administration; XML-RPC; Process management; Port scan, WMI;

I. INTRODUCTION

The world is driven by technology, in which the vast majority of companies, organizations and institutions rely on computers to some extent to connect their work. Whenever there is a problem with one or more computers within a network that requires system administration, and this cannot be delivered in a timely fashion, the delay in resuming the work process results in losses for the company or institution, losses that can be significant at times.

system administrators are in an increasing degree involved with the troubleshooting of solving many type of problems concerning the quality of service for the different applications[1].

According to [2] the common method to perform system administration is by accessing the remote system via network communication by means of client-server protocol.

Remote access via network communication are identified as ideal solution for performing system administration, irrespective of the administrator position. A fast, reliable and effective system administration services can be easily performed via remote connection.

The aim of this work is to develop an administration system capable of performing many administrating and controlling services remotely in a hidden manner. Therefore the remotely PC does not recognize any activity has performed by the administrator PC.

MPSA introduces new two advantages. The first one is implementing various important system administration and controlling remotely in an efficient and high speed performance, since the software depends on XML-RPC technique for managing the connection between server and the agents, in which the server performs services by calling specified procedure resides in the agents. The second advantage is introducing a portable client agent to work on different versions of Windows and Linux systems.

II. RELATED WORKS

Many efforts and application programs have been produced to facilitate the task of system administration. Machail, Craig and Janet presented NetReg program for remotely managing system registry for NT system[3]. This work can be used for specific system and perform limited system administration service.

According to Anis, Mohammad and Haissam, many different remote administration tools exists in the market, each of which offers some features for system administration[4]. These tools differs in its capabilities and platform they support, for example GoToMyPC is a powerful tool work on different platform except Windows7 and does not support services for managing essential system parts such as process, resource.

Sebastian and their partners suggested a new approach for system management services over a Wide Area Network (WAN) which performs easy selection and configuration of booting options for only Linux system[5].

Most of the previous works have used client-server approach to make a connection. This technique increases the network traffic loads. Moreover, the programs which have been introduced by previous work depends on client permission and works as foreground process.

MPSA has been implemented to overcome the above problems by using XML-RPC and Python libraries in different modern platform.

III. DIFFERENT TECHNIQUES WERE USED FOR REMOTE ADMINISTRATION:

This section presents different tools that had been used to administrate remote Linux and Windows systems and to access services for performing different tasks on the agents.

A. XML-RPC

As mentioned in [4] Remote Procedure Call (RPC) is a mechanism offers the capability of data exchange and invocation of functions residing in different process. That process can be on the same computer, on the local area network "LAN", or across the Internet. With RPC, essential program logic and related procedure code can exist on different computers, which is important for distributed applications.

In this work XML-RPC, which is a set of specifications that allow software running on disparate operating systems, have been used for running in different environments to make procedure calls over the network[6], since a heterogeneous environment in terms of operating systems are used.

B. WMI

Windows Management Instrumentation (WMI) is a set of Windows Driver model that provides an operating system interface, it allows scripting languages like Python to manage Microsoft Windows personal computers, both locally and remotely [7]. It is the management framework available in recent Windows systems. WMI is built on the COM "Component Object Model" infrastructure and can thus operate remotely, using DCOM "Distributed COM"[8].

WMI had been used in this work to access Windows system parts and information.

C. GTK

GTK (GIMP Toolkit) was originally developed as a toolkit for the GIMP (General Image Manipulation Program). It is a set of functions that have been used in implementation of screen snapshot in Linux and Windows systems[9].

D. Python Libraries:

Python has built-in support for the XML-RPC protocol and offers tools for implementing client-server applications without needing to install any additional packages.

In this work, python have been used to develop server and agents programs working in different operating system.

IV. SYSTEM ADMINISTRATION TASKS:

According to Eleen, the most important features/functions of the system administration are monitoring system activity, File management, system rebooting, and software monitoring [10]. This work implements an efficient and fast techniques to introduce a portable agent which contains all of these functions. The most important features which have been implemented by MPSA are described below:

The ability to inspect a running process and control its execution is a basic requirement security tool may require controlling opportunities. This work offering new capability for process administration. It uses an efficient method to view hidden processes depending on keep tacking of /proc system file, which contains information about each process. Also the work implements a new method for process deletion to delete specified process and all its children in order to avoid creation of zombie processes.

ii. System Resources Monitoring & Performance Controlling:

If system resources become too low, it can cause a lot of problems. The ability of resource monitoring can help to determine whether system is stable, or if some services need to be terminated or suspended temporarily depending on some criteria such as amount of CPU or memory usage.

iii. File Systems Monitoring

In any administration system, management files and directories represent an essential part. In this work, to manipulate files and directories, related system calls had been used in Linux and special API functions for Windows. The API functions and system calls provide complete control over the creation and maintenance of files and directories.

iv. Desktop Screenshots

A snapshot is the state of a system at a particular point in time. It can refer to an actual copy of the state of a system or to a capability provided by certain systems.

Implementing desktop snapshot remotely could provide a appropriate means for monitoring user activity in the target machine in any given time and rapidly.

v. Gathering System Information.

One of the basic task of system administration is how to find general system information when the system is running, such as CPU usage, the amount of memory on a system and its usage, and the amount of available disk space and its usage. Some of these tasks are performed repetitively, at regular intervals. Other tasks need to run only once.

vi. Port Scanning

This feature helps the administrator to check the network ports on the clients and to check the ports statues with giving the administrator the ability to close any unauthorized port to protect the clients' computers.

i. Process Administration

V. THE METHODOLOGY OF THE REMOTE SYSTEM ADMINISTRATION MODEL

MPSA consists of two main components: administration server and administration agent. The following scenario is used by software components to perform MPSA tasks: XML-RPC had been used to communicate between server and the agents program so that the agent programs defines functions including the implementation, the parameters and interfaces. These functions performing system information gathering and applied system tasks when the server calls one of the functions by following the function interface. The function may return a value contains system information performed into a suitable data structure. Figure (1) represents a high-level components of MPSA software.

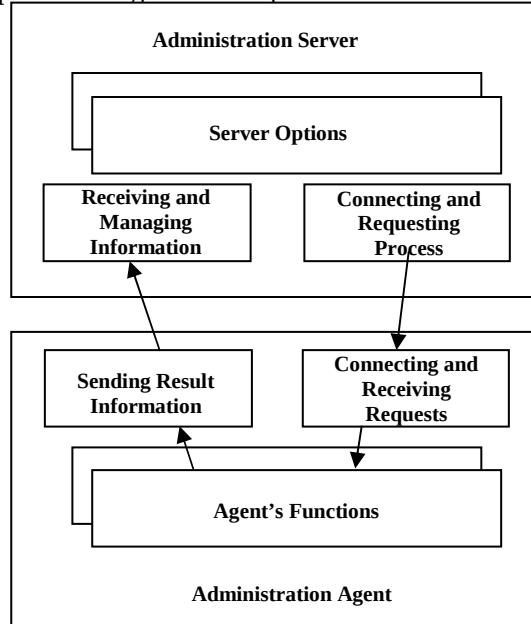


Figure (1) MPSA Main Components

a. Administration Server:

This part, which is responsible for controlling and managing all the clients agents, should be setup on the server computer. This part is responsible for detecting the opened network port among a range of IP addresses, which already had been specified for the clients computers in advance. When detecting an opened port for the target PC, the server performs checking operations to determine the type of operating system exists in the agent. The flowchart in figure (2) demonstrate the algorithm which is implemented by MPSA server program to administrate and control target system.

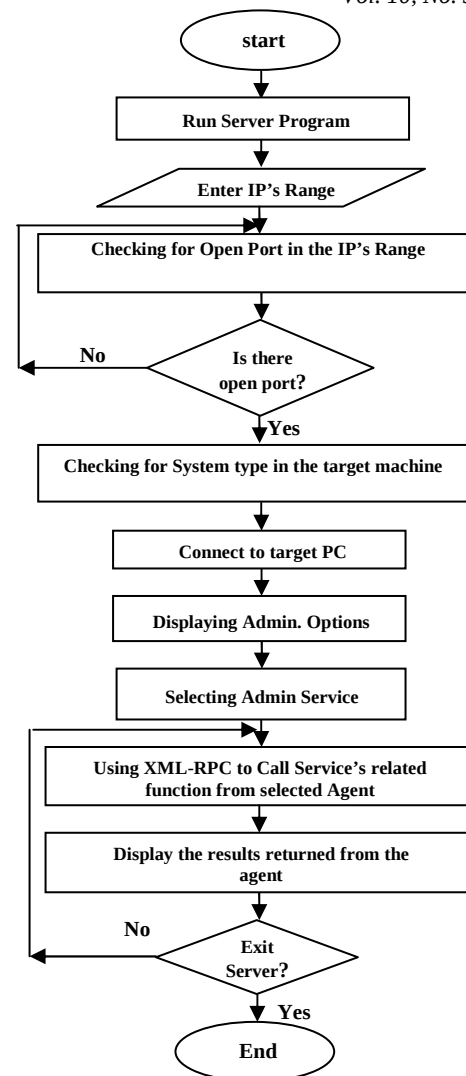


Figure (2) Administration Server

b. Administration Agent

MPSA's contains two agent programs to support system administration for different types of systems. These programs must be loaded by the administration for the first time, and be executed at system startup in Linux or Windows agent machines to make the agent system under administration. The server part will gather the information and pass operations and commands to be performed on the agent system using XML_RPC technique. On the other side agent program is responsible for performing the operations and passing back the resulted information to the server in the other side to be displayed. The flowchart in figure (3) demonstrate the algorithm which is implemented by MPSA agent program to perform tasks received from the server depending on agent system type.

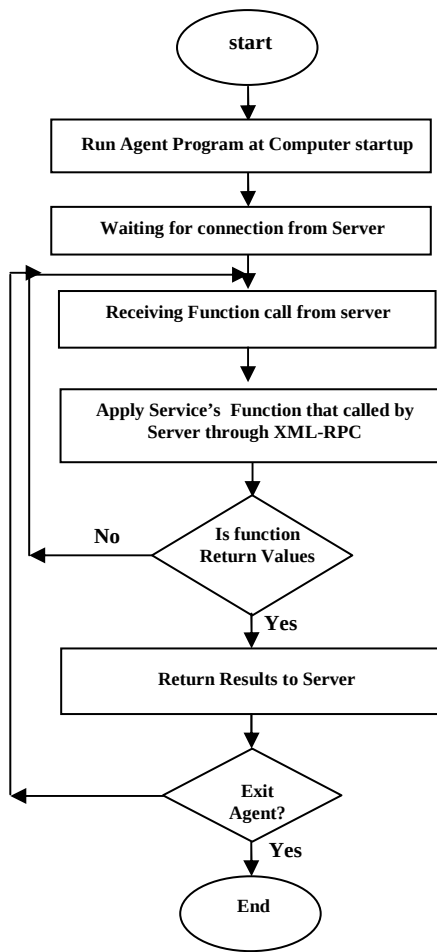


Figure (3) Adminstrating Agent

- The agent's programs for different system performs their task in a high speed time as a result of using XML-RPC protocol for responding for server requests.
- All tasks had been performed precisely on the target PC, and the system information and status are correct comparing to that obtained when using system tools such as task manager in Windows system or system manager in Linux system

Figure (4) depicts the first interface for MPSA server which is used to check for opened ports for the target computers and get information for their system then connect to the selected computer to start administration procedures using options explained in table (1).

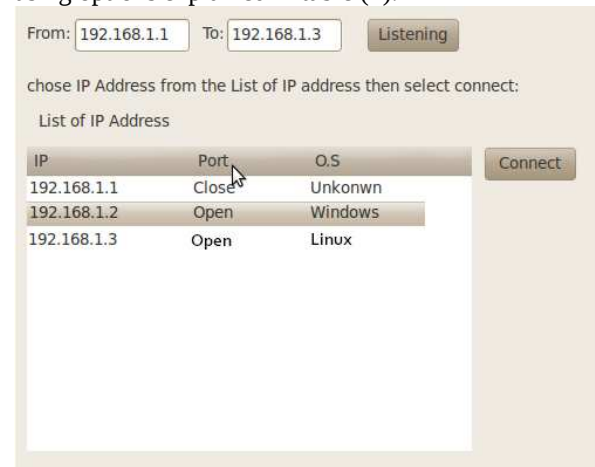


Figure (4) MPSA Server First Interface Window

Table (1) MPSA's Options

Main Option	Sub-Option	Function
Sys. Info.	CPU	Displays processor info.
	RAM	Displays memory info.
	OS	Displays OS info.
Process Admin.	Hidden Process	Shows all hidden process
	All Process	Shows all process
	Kill	Terminates specified process
	Suspend	Suspends specified process
	Resume	Resumes suspended process
Sys. Status	CPU	CPU Percentage usage

VI. EXPERIMENTS AND DISCUSSION

MPSA software implements remote system controlling and administration on different operating systems: Linux and Windows using Python language with different programming libraries such wxPython for performing software GUI, GTK, and WMI to manage many system parts in each of Linux and Windows. XML-RPC technique is used to exchange information between software parts.

MPSA software was tested successfully in the University of Mosul/Computer Sciences Dept. lab contains 5 computers. The first one was used as administration server working in Linux system with Ubuntu distribution, version 10.4. the others are used as agents working in Linux Ubuntu 10.4, Linux Ment12, Windows XP SP2, and Windows7. The results have shown efficiency in performance and speed in performing tasks on the target PC.

The most effective and powerful results are:

- It works efficiently on many types of systems such as: Windows with different versions (XP and 7) and Linux with various distributions and versions.
- Administration and controlling operations, that applied in the target PC did not appear any activity or be recognized from the agent user.

	RAM	RAM percentage usage
	Processes	No. of running, blocked processes
	Btime	time at which the system booted
	Disk Partitions	Disk partition usage & free
Packages		List of installed applications
Browser		File manager
Snapshot		get screen shot
Port Scan		List of open ports
Sys. Boot.	Shut Down	
	Restart	

The following figure represents administration server's options interface.

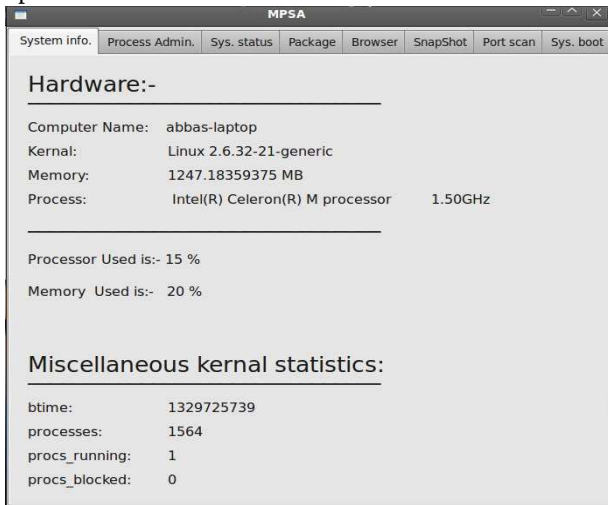


Figure (5) Administration Server's Options Interface

VII. CONCLUSIONS

This paper develops a new software for remote system administration and controlling different operating systems. The software depends on XML_RPC technique, which is fast and efficient method to exchange data and commands between server and agent. The software offers many valuable tools for controlling essential parts of target systems and gathering system information suitable for maintaining target system stability and controlling system usage, also system administration tools can be used to modify the behavior of system.

REFERENCES

- [1] Stig Jarle Fjeldbo, "Administration of Remote Computer Networks", Master Thesis, University of OSLO, Department of Informatics, 2005.
- [2] Marco Ramilli and Macro Prandini, "A messaging-based system for remote server administration", in proceeding of IEEE 3rd International Conference on Network and System security, 2009.
- [3] Michail Gomberg, Craig Stacey & Janet Sayre, "Scalable, Remote Administration of Windows NT", In Proceedings of the Second Large Installation Systems Administration of Windows NT Conference (LISANT), 1999.
- [4] Anis Ismail, Mohammed Hajjar and Haissam Hajjar, "Remote Administration Tools: A Comparative Study", Journal of Theoretical and Applied Information Technology (JATIT), 2008.
- [5] Sebastian Schmelzer, Dirk von Suchodoletz, Gerhard Schneider, Daniel Weingaertner, Luis Carlos E. de Bona and Carlos Carvalho, "Universal Remote Boot and Administration Service" in peocceeding of 7th Latin American Network Operations and Management Symposium Conference (LANOMS), 2011.
- [6] DI Siegfried G., xmlrpc-20020305, jugat, 2002.
- [7] Tim G., WMI v1.4.7 documentation, 2009, <http://timgolden.me.uk/python/wmi/contents.html>.
- [8] Microsoft Technet's Script Center, 2006. URL:<http://www.microsoft.com/technet/scriptcenter/>
- [9] Andrew Krause, "Foundation of GTK+ Development", APRESS, 2007.
- [10] Eleen Frisch, "Essential System Administration", 2002, 3rd Edit, O'Reilly.
- [11] Hanping Lufei, Weisong Shi and Vipin Chaudhary, "Adaptive Secure Access to Remote Services", in proceeding of IEEE International Conference on Services Computing, 2008.

An Efficient Automatic Attendance System Using Fingerprint Reconstruction Technique

Josphineleela.R

Research scholar

Department of Computer Science and Engineering

Sathyabamauniversity

Chennai,India

ilanleela@yahoo.com

Dr.M.Ramakrishnan

Professor/HOD-IT

Velammal Engineering College

Chennai,India

ramkrishod@gmail.com

Abstract— Biometric time and attendance system is one of the most successful applications of biometric technology. One of the main advantage of a biometric time and attendance system is it avoids "buddy-punching". Buddy punching was a major loophole which will be exploiting in the traditional time attendance systems. Fingerprint recognition is an established field today, but still identifying individual from a set of enrolled fingerprints is a time taking process. Most fingerprint-based biometric systems store the minutiae template of a user in the database. It has been traditionally assumed that the minutiae template of a user does not reveal any information about the original fingerprint. This belief has now been shown to be false; several algorithms have been proposed that can reconstruct fingerprint images from minutiae templates. In this paper, a novel fingerprint reconstruction algorithm is proposed to reconstruct the phase image, which is then converted into the grayscale image. The proposed reconstruction algorithm reconstructs the phase image from minutiae. The proposed reconstruction algorithm is used to automate the whole process of taking attendance, manually which is a laborious and troublesome work and waste a lot of time, with its managing and maintaining the records for a period of time is also a burdensome task. The proposed reconstruction algorithm has been evaluated with respect to the success rates of type-I attack (match the reconstructed fingerprint against the original fingerprint) and type-II attack (match the reconstructed fingerprint against different impressions of the original fingerprint) using a commercial fingerprint recognition system. Given the reconstructed image from our algorithm, we show that both types of attacks can be effectively launched against a fingerprint recognition system.

Keywords—Fingerprint Reconstruction, attendance management system, Minutiae Extraction

I. INTRODUCTION (HEADING 1)

Fingerprint reconstruction is one of the most well-known and publicized biometrics. Because of their uniqueness and consistency over time, fingerprints have been used for identification over a century, more recently becoming

automated due to advancements in computed capabilities. Fingerprint reconstruction is popular because of the inherent ease of acquisition, the numerous sources (e.g. ten fingers) available for collection, and their established use and collections by law enforcement and immigration.

Minutiae-based fingerprint matching algorithm [1] has been proposed to solve two problems: correspondence and similarity computation. For the correspondence problem, use an alignment-based greedy matching algorithm to establish the correspondences between minutiae.

Cryptographic techniques are being widely used for ensuring the secrecy and authenticity of information. Although several cryptosystems have proven security guarantees (e.g., AES and RSA), the security relies on the assumption that the cryptographic keys are known only to the legitimate user. Maintaining the secrecy of keys is one of the main challenges in practical cryptosystems. However, passwords can be easily lost, stolen, forgotten, or guessed using social engineering and dictionary attacks. Limitations of password-based authentication can be alleviated by using stronger authentication schemes, such as biometrics. Biometric systems establish the identity of a person based on his or her anatomical or behavioral traits, such as face, fingerprint, iris, voice, etc. Biometric authentication is more reliable than password-based authentication because biometric traits cannot be lost or forgotten and it is difficult to share or forge these traits. Hence, biometric systems offer a natural and reliable solution to the problem of user authentication in cryptosystems.

Reliable information security mechanisms are required to combat the rising magnitude of identity theft in our society. While cryptography is a powerful tool to achieve information security, one of the main challenges in cryptosystems is to maintain the secrecy of the cryptographic keys. Though biometric authentication can be used to ensure that only the legitimate user has access to the secret keys, a biometric system itself is vulnerable to a number of threats.

A critical issue in biometric systems is to protect the template of a user which is typically stored in a database or a smart card. The fuzzy vault construct is a biometric cryptosystem that secures both the secret key and the biometric template by binding them within a cryptographic framework. In [2], fuzzy vault scheme has been proposed based on fingerprint minutiae. Since the fuzzy vault stores only a transformed version of the template, aligning the query fingerprint with the template is a challenging task. We extract high curvature points derived from the fingerprint orientation field and use them as helper data to align the template and query minutiae. The helper data itself do not leak any information about the minutiae template, yet contain sufficient information to align the template and query fingerprints accurately. Further, we apply a minutiae matcher during decoding to account for nonlinear distortion and this leads to significant improvement in the genuine accept rate. The performance improvement can be achieved by using multiple fingerprint impressions during enrollment and verification.

Because of the stability and uniqueness, fingerprint is widely used in biometric identification. The matching method is one of the most crucial technologies in the Automated Fingerprint Identification System (AFIS). Whether two fingerprints are matched relies on the similarity measure between the effective features of them. There are mainly two kinds of features used in fingerprint matching: local features and global features. Two most prominent local ridge characteristics, called minutiae, are ridge ending and ridge bifurcation. Minutiae are the most widely used features in the matching process.

The performance of Automated Fingerprint Identification System (AFIS) is highly defined by the similarity of effective features in fingerprints. Minutia is one of the most widely used local features in fingerprint matching. In [3], proposes two global statistical features of fingerprint image, including the mean ridge width and the normalized quality estimation of the whole image, and proposed a novel fingerprint matching algorithm based on minutiae sets combined with the global statistical features. The algorithm proposed in this paper has the advantage of both local and global features in fingerprint matching. It can improve the accuracy of similarity measure without increasing of time and memory consuming.

The non-linear distortion in the fingerprint images makes it very difficult to handle matching as it changes the geometrical position of the minutiae points. The regions, that are affected, shift the geometry of the minutiae and hence pose a potential threat to acceptance of a genuine match. The distortion is due to the pressure applied on the scanner, the static friction, the skin moisture, elasticity, and rotational effects, which occur during the acquisition. The level of distortion increases from the center towards the outer regions. The existing approaches for fingerprint matching are: minutiae-based, and correlation-based. The former has several advantages over the latter such as lower time complexity, better space complexity, less requirement of hardware etc.

The uniqueness of a fingerprint is due to unique pattern shown by the locations of the minutiae points—irregularities of a fingerprint—ridge endings, and bifurcations. A novel minutiae-based approach [4], has been proposed to match fingerprint images using similar structures. Distortion poses serious threats through altered geometry, increases false minutiae, and hence makes it very difficult to find a perfect match. This algorithm divides fingerprint images into two concentric circular regions – inner and outer – based on the degree of distortion. The algorithm assigns weight ages for a minutiae-pair match based on the region in which the pair exists. The algorithm has two stages. In the first stage, the minutiae points are extracted, and in the second stage, the aligning and the matching of the fingerprint images are done. The algorithm is designed to reduce time taken in aligning, immediately after the calculation of the binary image.

Recent advances in automated fingerprint identification technology, coupled with the growing need for reliable person identification have resulted in an increased use of fingerprints in both government and civilian applications such as border control, employment background checks, and secure facility access. In [5], Quadratic differentials naturally define analytic orientation fields on planar surfaces. This method proposed model orientation fields of fingerprints by specifying quadratic differentials which is used for reliable person identification. Models for all fingerprint classes such as arches, loops and whorls are laid out. These models are parameterized by few, geometrically interpretable parameters which are invariant under Euclidean motions. Potential applications of these models are the use of their parameters as indices of large fingerprint databases, as well as the definition of intrinsic coordinates for single fingerprint images. The accuracy of models is still challenging task for arches.

General characteristics of the fingerprint emerge as the skin on the fingertip begins to differentiate. Fingerprint recognition systems have the advantages of both ease of use and low cost. Because among various biometric identifiers, such as face, signature, and voice, the fingerprint has one of the highest levels of distinctiveness and performance and it is the most commonly used biometric modality. Haiyun Xu et. al., [6], proposed a novel method to represent minutiae set as a fixed-length feature vector, which is invariant to translation, and in which rotation and scaling become translations, so that they can be easily compensated for recognition. These characteristics enable the combination of fingerprint recognition systems with template protection schemes that require a fixed-length feature vector. This method introduces the concept of algorithms for two representation methods: the location-based spectral minutiae representation and the orientation-based spectral minutiae representation. Both algorithms are evaluated using two correlation-based spectral minutiae matching algorithms. The performance can be improved by using a fusion scheme and singular points. The spectral minutiae representation overcomes the drawbacks of the minutiae sets, thus broadening the application of minutiae-based algorithms. The minutiae extractor is not reliable it affects the efficiency of spectral minutiae representation.

Automated Fingerprint Identification Systems (AFISs) have played an important role in many forensics and civilian applications. There are two main types of searches in forensics AFIS: ten print search and latent search. In ten print search, the rolled or plain fingerprints of the 10 fingers of a subject are searched against the fingerprint database of known persons. In latent search, a latent print developed from a crime scene is searched against the fingerprint database of known persons. Latent Fingerprint matching [7], propose a system for matching latent fingerprints found at crime scenes to rolled fingerprints enrolled in law enforcement databases which overcomes the difficulties in poor quality of ridge impressions, small finger area, and large nonlinear distortion. In addition to minutiae, extended features are also used including singularity, ridge quality map, ridge flow map, ridge wavelength map, and skeleton. In order to evaluate the relative importance of each extended feature, these features were incrementally used in the order of their cost in marking by latent experts. The matching accuracy should be improved.

Despite tremendous progress made in automatic fingerprint identification systems, matching incomplete or partial fingerprints such as latent prints remains a critical challenge today. Existing partial fingerprint algorithms concentrate on improving the accuracy of one-to-one matching based on local ridge details. However, the performance of one-to-one matching depends on image quality as well as the number of high-level features detectable in the partial fingerprint segments. These ad hoc algorithms are designed on the basis of more delicate one-to-one comparisons. When used in one-to-many applications, they generally assume sequential matching or that the candidate list for such matching has already been established. However, sequential matching is not efficient for large-scale identification, which can involve thousands or millions of records in the target database, and retrieving a short and reliable list of candidates for matching is difficult in practice. An innovative method [8], propose an analytical approach for reconstructing the global topology representation from a partial fingerprint. Analytical approach solves the problem of retrieving candidate lists for matching partial fingerprints by exploiting global topological features. First, an inverse orientation model for describing the reconstruction problem is presented. Then, a general expression for all valid solutions to the inverse model is provided. This allows us to preserve data fidelity in the existing segments while exploring missing structures in the unknown parts. Further developed algorithms for estimating the missing orientation structures based on some a priori knowledge of ridge topology features are described. The statistical experiments show that the proposed model-based approach can effectively reduce the number of candidates for pair wise fingerprint matching, and thus significantly improve the system retrieval performance for partial fingerprint identification.

Fingerprint matching systems generally use four types of representation schemes: grayscale image, phase image, skeleton image, and minutiae, among which minutiae-based representation is the most widely adopted one. It has

been traditionally assumed that minutiae template does not retrieve any information about original fingerprint. In [9], three levels of information about the parent fingerprint can be elicited from a given minutiae template: the orientation field, the fingerprint class, and the friction ridge structure. The orientation estimation algorithm determines the direction of local ridges using the evidence of minutiae triplets. The estimated orientation field, along with the given minutiae distribution, is then used to predict the class of the fingerprint. Finally, the ridge structure of the parent fingerprint is generated using streamlines that are based on the estimated orientation field. Line Integral Convolution is used to impart texture to the ensuing ridges, resulting in a ridge map resembling the parent fingerprint. But the visual appearance of reconstructed fingerprint is not accurate.

The location, position, as well as the type and quality of the "minutiae" are factors taken into consideration in the template creation stage. A minutiae-based template did not contain enough information to allow the reconstruction of the original fingerprint. A novel approach [10], is proposed to reconstruct fingerprint images from standard templates and examines to what extent the reconstructed images are similar to the original ones. The efficacy of the reconstruction technique has been assessed by estimating the success chances of a masquerade attack against nine different fingerprint recognition algorithms. The experimental results show that the reconstructed images are very realistic and that, although it is unlikely that they can fool a human expert, there is a high chance to deceive state-of-the-art commercial fingerprint recognition systems.

The fingerprint recognition system is used for person authentication and identification in industries and many commercial appliances. The fingerprint recognition does not have the efficiency in the case of fake fingerprints which extracts minutiae from templates. The compactness of minutiae representation has created an impression that the minutiae template does not contain sufficient information to allow the reconstruction of the original grayscale fingerprint image. In [11], a novel fingerprint reconstruction algorithm is proposed to reconstruct the phase image, which is then converted into the grayscale image. Reconstruction algorithm not only gives the whole fingerprint, but the reconstructed fingerprint contains very few spurious minutiae. A fingerprint image is represented as a phase image which consists of the continuous phase and the spiral phase. The proposed reconstruction algorithm has been evaluated with respect to the success rates of type-I attack and Type II attacks using a commercial fingerprint recognition system. Reconstruction algorithm should be modified in order to apply the important problems of latent fingerprint restoration. The proposed reconstruction algorithm is used to automate the whole process of taking attendance, manually which is a laborious and troublesome work and waste a lot of time, with its managing and maintaining the records for a period of time is also a burdensome task.

II. ATTENDANCE MANAGEMENT SYSTEM

Attendance management system is one of the most successful applications of biometric technology. With the integration and use of biometric technology getting simpler, many institutions are venturing down the biometric road to verify the time and attendance of their students and staffs.

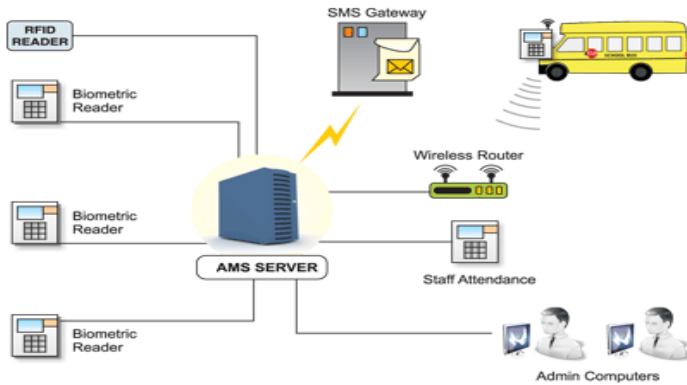


Figure 1. Attendance Management System (AMS)

III. Related work

In this system the fingerprint is taken as an input for attendance management and it is organized into the following modules Pre-processing, Minutiae Extraction, Reconstruction, Fingerprint Recognition, Report generation

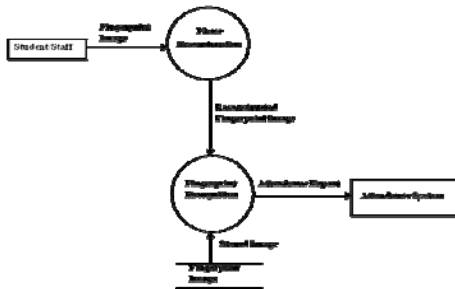
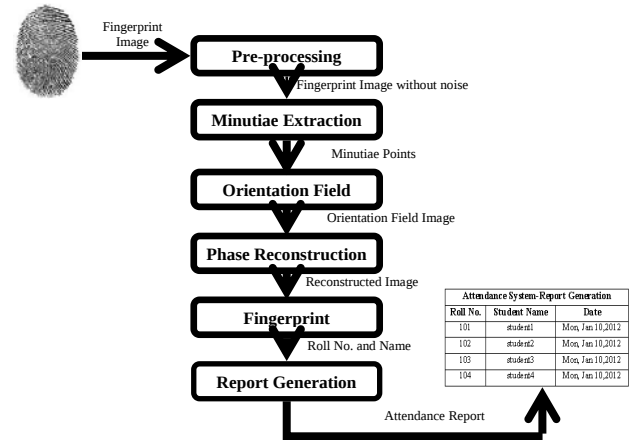


Fig 2.DFD for Attendance Management System

SYSTEM ARCHITECTURE



IV. Preprocessing

There are two steps in Pre-processing

Step 1: Segmentation

Step 2: Normalization

1 Segmentation

- Image segmentation separates the foreground regions and the background regions in the fingerprint image.
- Segmentation is a process by which can discard these background regions, which results in more reliable extraction of minutiae points.

2 Normalization

- Normalization is a process of standardizing the intensity values in an image so that these intensity values lie within a certain desired range.
- It can be done by adjusting the range of grey-level values in the image.

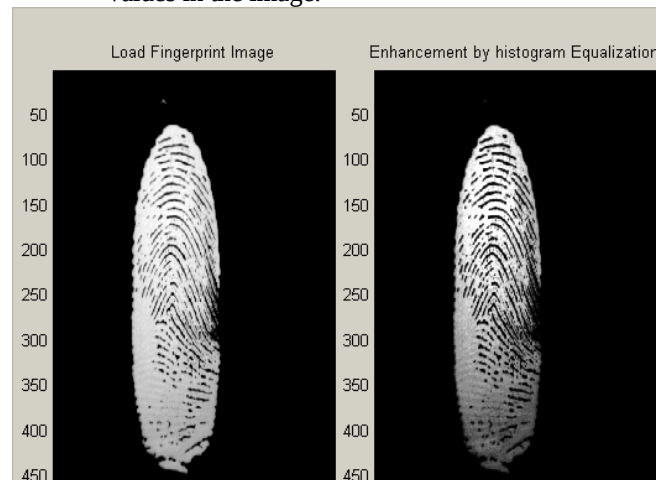


Fig 3 Result of histogram equalization (a) original image (b) After histogram equalization

V. Minutiae Extraction

Minutiae points are extracted from composite phase image of fingerprint image which is obtained by adding spiral phase to the continuous phase.

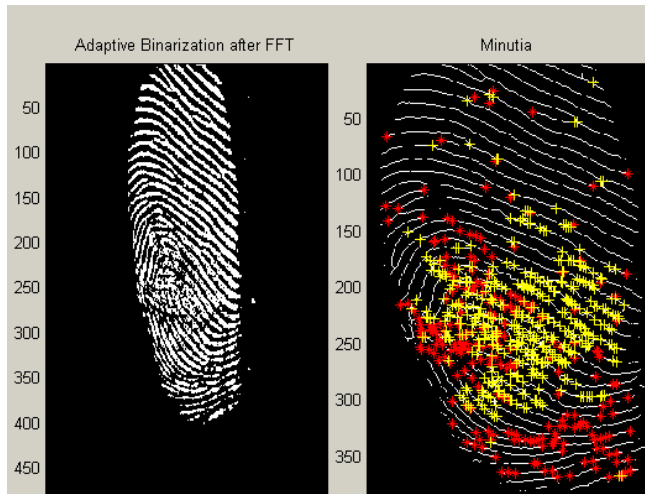


Fig 4 Result of Minutiae Extraction

VI. Fingerprint Reconstruction

There are two steps in reconstruction

Orientation Field Reconstruction

- An orientation field reconstruction algorithm that can work even when only one minutia is available.
- The image is divided into non overlapping blocks of 8×8 pixels and an orientation value is computed for each foreground block.

Phase Reconstruction

- The continuous phase has been reconstructed at all of the foreground blocks by estimating the phase offset value.
- The reconstructed phase image validates the minutiae points and eliminates spurious minutiae.

VII. Fingerprint Recognition

Fingerprint is recognized if the reconstructed fingerprint matches with the original fingerprint.

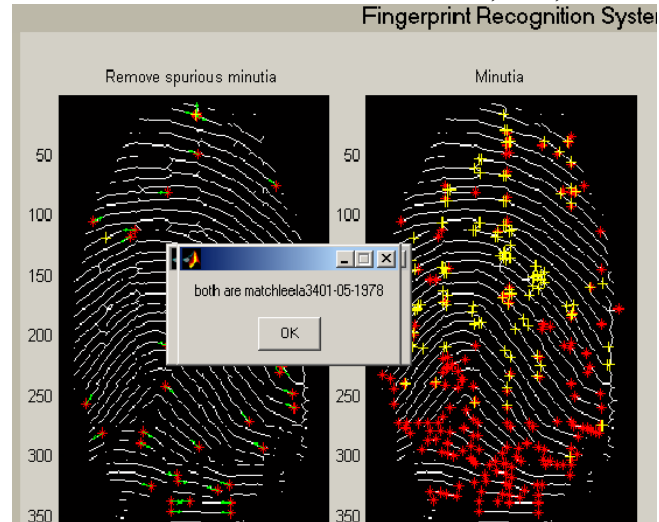
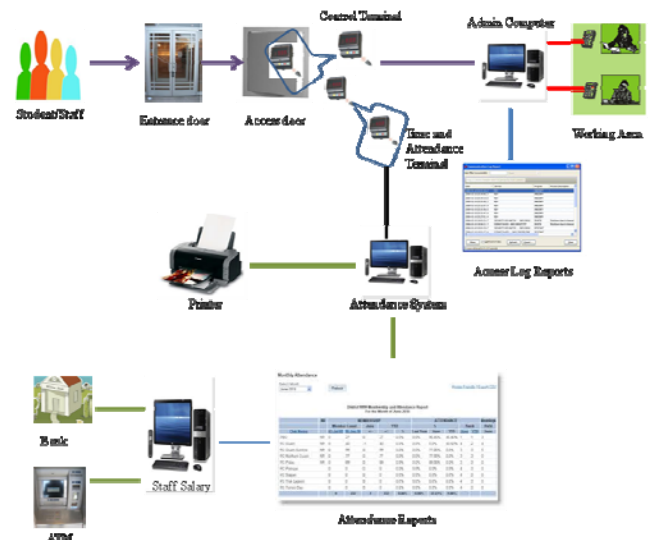


Fig 5 Result of Minutiae Matching

VIII. Result

The report will be generated with Roll number of the matched fingerprint and stored in an attendance system.



IX. Conclusion

The proposed system will make way for perfect management of students and staff attendance and produce more accuracy

REFERENCES

- [1] Jianjiang Feng, "Combining minutiae descriptors for fingerprint matching", Pattern Recognition, pp. 342 – 352, April 2007.
- [2] Karthik Nandakumar, Anil K. Jain, and Sharath Pankanti, "Fingerprint-Based Fuzzy Vault: Implementation and Performance", IEEE transactions, Vol. 2, No. 4, December 2007.
- [3] Peng Shi, Jie Tian, Qi Su, and Xin Yang, "A Novel Fingerprint Matching Algorithm Based on Minutiae and Global Statistical Features", IEEE Conference, 2007.
- [4] Neeta Nain, Deepak B M, Dinesh Kumar, Manisha Baswal, and Biju Gautham "Optimized Minutiae-Based Fingerprint Matching", Proceedings, 2008.
- [5] Stephan huckemann, Thonmas Hotz, and Axel Munk, "Global Models for the Orientation Field of Fingerprints: An Approach Based on Quadratic Differentials", IEEE Transactions, Sep. 2008.
- [6] Haiyun Xu, Raymond N.J. Veldhuis, Asker M. Bazen, Tom A.M. Kevenaar, Ton A.H.M. Akkermans, and Berk Gokberk, "Fingerprint Verification Using Spectral Minutiae Representations", IEEE Transactions, Vol. 4, pp. 397-409, Sep. 2009.
- [7] Anil K. Jain, and Jianjiang feng, "Latent Fingerprint Matching", IEEE Transactions, Vol. 33, pp. 88-100, Jan. 2011.
- [8] Yi (Alice) Wang, and Jiankun Hu, "Global Ridge Orientation modelling for Partial Fingerprint Identification," IEEE Transactions, Vol. 33, Pg. 72-87, Jan. 2011.
- [9] A. Ross, J. Shah, and A.K. Jain, "From Template to Images Reconstructing Fingerprints from Minutiae Points", IEEE Transactions, Vol. 29, pp. 544-560, Apr. 2007.
- [10] R. Cappelli, A. Lumini, D. Maio, and D. Maltoni, "Fingerprint Image Reconstruction from Standard Templates", IEEE Transactions, Vol. 29, pp. 1489-1503, Sept. 2007.

- [11] Jianjiang Feng, and Anil K. Jain, "Fingerprint Reconstruction: From Minutiae to Phase", IEEE Transactions, Vol. 33, No. 2, pp. 209-223, Feb. 2011.

AUTHORS PROFILE

Authors Profile ...

R. Josphineleela received the B.Sc computer science degree from the Department of Computer Science Madurai Kamaraj University, Madurai, India, in 1998 and M.C.A degree in the same university in 2001. She received the M.E. degree from the Department of Computer Science and Engineering Sathyabama University, Chennai, India, in 2007. She has published Four papers in International Level Conferences and Three papers in National level Conferences. She has published one paper in International Journal. She has 10 years teaching experience and was awarded best teacher in the year 2011 by Panimalar Engineering College, Chennai. She is pursuing her PhD under the guidance of Dr. M. Ramakrishnan. Her research interests are in Image processing, Pattern recognition, soft computing and artificial neural network etc.

Dr. M. Ramakrishnan was born in 1967. He is working as a Professor and Head of PG department of Computer Science and Engineering in velammal Engineering College, Chennai. He is a guide for research scholars in many universities. His area of interest is Parallel Computing, Image Processing, Web Services and Network Security. He has 21 years of teaching experience and published 8 National and International journals and 40 National and International Conferences. He is member of ISTE and senior member of IACSIT.

Towards More Realistic Mobility Model in Vehicular Ad Hoc Network

Dhananjay S. Gaikwad, MTech,
Assistant professor, ghananjayg63@gmail.com

Mahesh Lagad, MTech(Appear)
Assistant professor, maheshlagad@gmail.com

Prashant Suryawanshi, MTech(Appear)
Assistant professor, sprashant1234@gmail.com.

Vaibhav Maske, MTech(Appear)
Assistant professor, vamaske@gmail.com

Computer Engineering Department,
HSBPVT'S GOI, College of Engineering,
Kashti. India- 414701.

Abstract— Mobility models or the movement patterns of nodes communicating wirelessly, play a vital role in the simulation-based evaluation of vehicular Ad Hoc Networks (VANETs). Even though recent research has developed models that better corresponds to real world mobility, we still have a limited understanding of the level of the required level of mobility details for modeling and simulating VANETs. In this paper, we propose a new mobility model for VANETs that works on the city area and map the topology of streets and behavior of vehicles at the intersection of roads. Our model change the speed of nodes after some specific distance in accordance to neighboring nodes that is according to a density of nodes, so that this will lead to a realistic situation on the roads. Our model accounts the various characteristics of VANETs such as traffic lights, acceleration/deceleration due to nearby vehicles, attraction points where maximum numbers of vehicle tends to go. Using the real and controlled map of street, we compare our mobility model with the random direction mobility model. Our result demonstrates that probability of link availability in VANETs is more sensitive to the vehicles waiting at intersections and acceleration/deceleration of vehicles. We also found that probability of link availability suffers at the intersection of the roads; because of some nodes cross the signal continue movement in horizontal direction while some nodes change the direction of traveling to vertical.

Keywords- Vehicular Ad Hoc Network (VANETs), Mobile Ad Hoc Network (MANETs).

I. INTRODUCTION

Ad hoc network is a collection of wireless mobile nodes without any fixed base station infrastructure and centralized management. Each node acts as both host and router, which moves arbitrarily and communicates with each other via multiple wireless links. It is a multi-hop wireless network, where

packets need to pass through several nodes to reach destination [2]. Vehicular Ad Hoc Network (VANETs) are a special case of Mobile Ad Hoc Network (MANETs) and consist of a number of vehicles traveling on urban streets capable of communicating with each other without fixed infrastructures. VANETs are expected to benefit safety applications, gathering and disseminating real-time traffic congestion and routing information, information services Such as transparent connection to internet etc [3].

One critical aspect of VANETs simulation is the movement pattern of vehicles, also called mobility models. Mobility model determine the location of nodes in the topology at any given instant, which strongly affects network connectivity and throughput [5]. There are several mobility models such as random pattern, graph constrained commonly used in popular wireless simulators such as ns2 [16] by VANET researchers [4]. But one problem with these models is that they ignores some critical aspects of the real world traffic such as queuing of vehicles at road intersection, traffic lights and traffic signs, acceleration and deceleration according to neighbor vehicles. Mobility models should reflect as possible as the real behavior of vehicular traffic on the road [1]. In this paper, we propose a new mobility model, which incorporates important features of mobility model on the road, such as presence of traffic light on the road, node movement is restricted to the road structure and speed changes in accordance to the neighboring vehicle.

The rest of paper is organized as follows: section 2 describes some currently used mobility models and some tools for generation of mobility models. Section 3, describes our proposed mobility model and simulation of our model. Finally Section 4

concludes the paper.

II. MOBILITY MODEL OVERVIEW

Mobility model reflects the behavior of the nodes throughout the simulation time. It shows how the nodes change their speed and direction in accountancy to the neighboring vehicles and according to the traffic rule. Following are some important factors those affect the mobility of nodes in VANETs.

- *Street structure*: Streets force nodes to confine their movements to well-defined paths. This constrained movement pattern determines the spatial distribution of nodes and their connectivity. Streets can have either single or multiple lanes and can allow either one-way or two-way traffic.
- *Block size*: A city block can be considered the smallest area surrounded by streets. The block size determines the number of intersections in the area, which in turn determines the frequency with which a vehicle stops.
- *Traffic control mechanisms*: The most common traffic control mechanisms at intersections are stop signs and traffic lights. These mechanisms result in the formation of clusters and queues of vehicles at intersections and subsequent reduction of their average speed of movement.
- *Interdependent vehicular motion*: Movement of every vehicle is influenced by the movement pattern of its surrounding vehicles.
- *Average speed*: The speed of the vehicle determines how quickly its position changes, which in turn determines the rate of network topology changes.

Whenever a mobility model is designed for VANETs, that model should consider the factors that affect the mobility of node. These types based on the number of nodes considered while designing the model and the way mobility information is stored.

2.1 Trace based mobility model

This type of model [6] is suitable to emulate the real scenarios in MANET and VANET. Traces describe the movement of vehicles throughout the simulation. Traces are the best information to find the mobility patterns of node, if we have traces of long period and involvement of many participants. Traces reflect the movement histories of the nodes in the network. We can expect mobility patterns provided by them lead to realistic mobility modeling. But the VANET applications are not widely deployed; there are fewer traces for evaluation. Another issue related to traces is, the nature of network is decentralized and

difficult to collect the real time traces of all the nodes. Following section describe mobility models that generate the trace file, contain the traces of vehicles movement.

2.1.1 Mobility model generator for Vehicular Network (MOVE): Mobility model generator for Vehicular Network (MOVE) [12] facilitates users to rapidly generate realistic mobility models for VANET simulation with a visualization property. This model works with another micro-simulator traffic model, called SUMO [13]. MOVE model consists of two main components: Map Editor and Vehicle Movement Editor. Map Editor is used to create the road topology, which is either created by manually, automatically or by importing the maps from databases such as TIGER ((Topologically Integrated Geographic Encoding and Referencing). The Vehicle Movement Editor used for the generation of vehicle movement. The output of MOVE is a mobility trace file which contains the information on vehicle movement that can be used by network simulator. All the parameter configuration of vehicle movement is done in a static way. This model does not consider micro – mobility features.

2.1.2 Street Random Waypoint (STRAW): Street Random Way point (STRAW) is a tool [14] that generates the mobility patterns with extraction of urban topologies from the TIGER database. It supports for the micro – mobility features of models. STRAW implements a complex intersection management using traffic lights and traffic signs. Due to this characteristic, vehicle shows a more realistic behavior when reaching at intersection. It includes a traffic control mechanisms that force drivers to follow deterministic admission control protocol when encountering intersection. Drawback of STRAW model is it does not give details about the traffic flows. Also it does not specify the lane changing behavior.

2.2 Entity mobility model

Entity mobility model represents mobile node as a random entity which moves randomly over the observed area, where speed and directions of node independent with the neighboring nodes.

2.2.1 Random Walk mobility mode: Random Walk mobility model is Entity mobility model, in which mobile node moves from its current location to a new location by randomly choosing a direction and speed in which to travel [4]. The new speed and direction are both chosen from pre-defined ranges, respectively [min-speed,max-speed] and $[0,2\pi]$ respectively. Each movement in the Random Walk Mobility Model occurs in either a constant time interval T or a constant traveled distance, at the end of which a new direction and speed are calculated.

2.2.2 Random way point model: In Random way point model [7], mobile nodes move randomly and freely

without any restrictions. In this model, the destination, speed and direction all are chosen randomly and independent of other nodes. The fraction of nodes in network remains static for the entire simulation time. The velocity of node is uniformly chosen at random from the interval $[V_{min}, V_{max}]$. The node moves towards destination with a velocity v . When it reaches to destination, it remains static for the predefined pause time and moving again according to same rule. The mobility behavior of nodes very much depends on the pause time and maximum speed of nodes. The parameters to describe a simulation setup of model are Size and shape of the deployment region Q , initial spatial node distribution $f_{int}(x)$, static parameter p_s , with $0 < p_s < 1$, Probability density function $f_{Tp}(tp)$ of pause time, Minimum and maximum speed : $0 < V_{min} \leq V_{max}$.

The components of node distribution $f(x)$ is composed of three distinct components as shown in equation. $(f(x) = f_s(x) + f_p(x) + f_m(x) \dots (1))$

2.2.3 Gauss Markov model: Random way-point model generates speed and direction of nodes independent on previous history. It directly selects speed and direction from its predefined range, so this can create a sudden stop and sharp turn problem. Gauss Markov model [8] first calculates the speed and direction of movement for each node. Then nodes move with the calculated speed and direction for a period. After that period similar movements begins again. The time that is used in the movement in each interval before the change in speed and direction, is constant. The current speed and direction related to the previous speed and direction shown by equation (2) and (3).

$$\begin{aligned} \square \quad s_n &= \alpha_{s_{n-1}} + (1-\alpha) * s + (1-\alpha^2) * s_{x_{n-1}}^{1/2} \\ \square \quad d_n &= \alpha_{d_{n-1}} + (1-\alpha) * d + (1-\alpha^2) * s_{d_{n-1}}^{1/2} \end{aligned}$$

Where, s_n and d_n are the values of speed and direction about its individual reference point. α is the constant value in the range $[0, 1]$, s and d are constants representing the mean speed and direction, $\alpha_{s_{n-1}}$ and $\alpha_{d_{n-1}}$ are variables from a Gaussian distribution. Gauss Markov model overcomes sudden stop and sharp turn problems of Random way point model [7].

2.2.4 Random direction mobility model: Random direction mobility model is, besides the random waypoint model, probably the most widely used model. This model considers individuals moving on straight walk segments with constant speed and optional pauses between the walk segments. In this model each node alternates periods of movement (move phase) to periods during which it pauses (pause phase) [10]; at the beginning of each move

phase, node independently selects its new direction and speed of movement. Speed and direction are kept constant for the whole duration of node move phase.

2.3 Group Mobility model

Entity mobility models represent multiple mobile nodes whose actions are completely independent of each other. In an ad hoc network, however, there are many situations where it is necessary to model the behavior of mobile nodes as they move together. For example, a group of soldiers in a military scenario may be assigned the task of searching a particular plot of land in order to destroy land mines, capture enemy attackers. In order to model such situations, a group mobility model is needed to simulate this cooperative characteristic. In this section, we present reference group mobility models.

Reference Point Group Mobility model: The Reference Point Group Mobility model represents the random motion of a group of mobile nodes as well as the random motion of each individual mobile node within the group. Group movements are based upon the path traveled by a logical center for the group. It is used to calculate group motion via a group motion vector, GM. The motion of the group center completely characterizes the movement of this corresponding group of mobile nodes, including their direction and speed. Individual mobile nodes randomly move about their own pre-defined reference points whose movements depend on the group movement [11]. As the individual reference point move from time t to $t+1$, their locations are updated according to the group's logical center. Once the updated reference group points, RP ($t+1$) are calculated, they are combined with a random vector, RM, to represent the random motion of each mobile node about its individual reference point. The length of RM is uniform distributed within a specified radius centered at RP ($t+1$) and its direction is uniformly distributed between 0 and P_i .

3. PROPOSED MOBILITY MODEL

3.1 Our proposed model

We develop our model by considering the real scenarios on the roads. We consider several parameters of real traffic situation on the roads of city. The parameters attraction point, speed variation, traffic light are considered. We change the speed of nodes after some specific distance according to the density of the nodes. We model the signal based on the horizontal traffic time stamps and vertical traffic time stamps.

Attraction point: In real scenarios, vehicles do not move randomly from one point to another. They used to set some fixed destination to reach. We consider this fact in our proposed model. Vehicles generally move in deterministic way. Whenever vehicles enter in the city, they not always move in straight line. They may change their direction of movement to a specific point according to the importance of that point. Suppose if the vehicle rider is a student, then he will definitely move to the college road. Like this if vehicle rider want to go restaurant, he will move to the restaurant road. Users move in group's towards the attraction points. More number of nodes will be around the most attractive point compare to less attractive.

In our model we have modeled the attraction point as the function of probability value at signaling point. In our model, each vehicle has assigned some probability values whenever they enter in the simulation area according to their attraction behavior. The checking of the probability value is done at the signaling point.

Speed variation: Generally whenever vehicles enter in the simulation area as city area in our proposed model, their speed do not remains constant throughout the simulation time and area. The speed of vehicle is changes in according to the neighboring vehicles, traffic lights, street layout, and pedestrian movement.

In our proposed model, we change the speed of the vehicle after some specific distance. So it gives a scenarios that a vehicle changes its speed in according to the neighboring vehicles. So when any vehicle comes closer to already moving vehicle for the overtaking, the moving vehicle will decrease the speed for some time and overtaking node will increase the speed at that time. This speed variation parameter will increase the realism of our proposed model.

Traffic light: Traffic light is for the management of intersection on the roads. In our proposed model, we have modeled the traffic light as a coordinated traffic light. For that first consider the single horizontal lane. The light turns green in such a manner that only traffic along the single lane cross the intersection simultaneously. Vehicles that need to turn left will go directly. We have modeled the signal in such that when the traffic light turns red, the vehicles that need to cross the signal will wait other vehicle directly take turn.

Simulation area: In our proposed model, we have considered the area of city for the deployment of our model. In that vehicles enter the simulation area on the left side of the road. A vehicle then moves towards the right direction with speed range from 5m/s to 25m/s. There are traffic lights at the

intersection point, to handle the traffic in both the direction.

Node movement: At the beginning, the nodes are distributed along the starting points of the horizontal and vertical lane. Nodes are allowed to move only through the predefined paths. Node knows the distance from the origination point or arriving point of traffic to the signaling point. The nodes have assigned probability values according to their point of interest. Each node calculates the time required to reach at signal. The traffic signal is modeled on the time stamp basis. At the signal, each node checks its next direction of movement according to the probability value and attraction point, i.e. destination point.

3.2 Operation of proposed mobility model

In this section, we describe algorithm of movement of mobile nodes of proposed mobility model. Initially, all nodes start from the initial point on the road and moves up to the boundary of the simulation area.

1. Define all the necessary variables signal time, signal coordinates, number of nodes, area of simulation, TURN.
2. Do for all nodes, from 1 to n, where n is number of nodes in the simulation.
3. Node movement starts from initial point towards the boundary of simulation area in both directions horizontal and vertical.
4. Set distance after which the node speed will change to 15m and speed of node= random value between {5m/s and 25m/s}.
 - a. Calculate the time required to travel given dist by the formula, $\text{time} = \text{dist} / \text{speed}$.
 - b. Add time and dist to the total time and total distance of the node respectively.
5. If $\text{node_distance} == \text{signal}$, then
Check the time of signal and the probability value of node.
If $\text{signal_time} == \text{node_time}$ (that means signal is red) and probability value! = TURN. Then wait for the remaining time till the signal turns green and go to step 8.
Otherwise, cross the signal without waiting. And take the turn and go in upward direction without stopping at signal and go to step 8.
6. Continues the nodes movement and go to step 5.
7. While (end of simulation area).

8. Take a movement for other nodes and go to step 2.
9. While (number of nodes)
10. End
11. Now distance matrix contains the time and distance values for nodes.
12. Take the values from the distance matrix.
13. Calculate the probability of link availability and average number of neighbors per node.
14. Compare these values with random waypoint, gauss Markov, random direction and city section models.
15. Show the results.

3.3 Comparison of mobility model

Table 1

Features	Random waypoint model	Our model
Horizontal and vertical lane	No horizontal and vertical lane, nodes are move randomly around simulation area.	Horizontal and vertical lane present, which shows the road of the city.
Cross point	No cross point	Cross points are present, show the road intersection.
Attraction point	No attraction point	Check the probability of the nodes.
Speed variation	Constant for some time interval	Change after some specific distance.

Table 1 shows broadly the comparison between our proposed mobility model and random waypoint model.

3.5 Simulation results

We performed simulation using the MATLAB. All nodes position is shown by their x and y coordinate values. We are not taking into account the third dimension (z- direction) of position. So nodes are assumed to move in two dimensional planes all the time. All nodes initialized by their initial position and make them travel to a specified destination point. Simulation parameters are shown in the table 2.

A. Varying number of nodes

This section compares the mobility models with different number of nodes in 800mX800 road topology. Figure 1 compares our proposed mobility model with random direction models with 20 numbers of nodes.

The results indicate that probability of link availability is higher at the initial time of simulation and it is gradually decreases as time passes i.e. as

node moves towards the end of simulation area and time for the random direction mobility model. This is because of each node takes random direction independent on the previous node and that direction is not restricted to the particular road topology. But in our proposed model, the probability of link availability decreases only at the signaling point and probability value remains high for rest time of simulation. This

Table 2

Parameter	Values
Simulation time	100 seconds
Number of nodes	10,20,30,40,50,75
MATLAB version	MATLAB 7.8.0.347
Mobility model characteristics	Horizontal and vertical lanes.
popularity of attraction points	Randomly assigned between 0 to 1.
MAXSPEED	25 m/s
MINSPEED	5 m/s.
Distance after that speed changes	10m, 15m.
Mobility model	Our proposed model and Random direction model.

would give the actual situation on the roads.

When we increase the number of nodes from 20 to 30, the figure 2 shows the results. The results indicate that, there is improvement in the probability of link availability. This is due to the increase in the density of nodes.

B. Behavior of our mobility model with varying number of nodes.

Case 1: This section shows the behavior of our proposed mobility model, by varying the number of nodes, while keeping the speed constant between 5m/s to 25m/s. The figure 3 shows the results. The results indicate that the probability value increases with the increase in the number of nodes in the simulation. This is because the density of nodes increases number of connections between the nodes.

Case 2: This section shows the behavior of our model with varying number of nodes and increasing the speed of nodes. The figure 4 shows the results with the increase in the speed of nodes. The speed of nodes is in between 0m/s to 25m/s. The result indicates that there is improvement in probability of link availability value for 10 and 20 nodes. But the probability values decrease for 30 and 40 nodes and again increases for 50 nodes. So by comparing results

in figure 3 and figure 4, we can show that the speed of the nodes play an important role during the communication. If we increase the speed of nodes that does not mean that we get the better results. So the performance of routing protocols very much depends on the speed and density of the nodes.

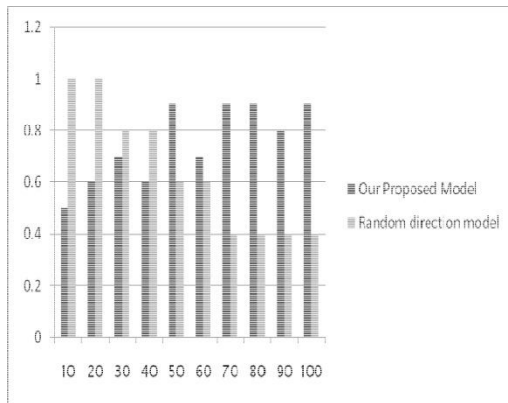


Figure 1: Probability of Link Availability versus Time.

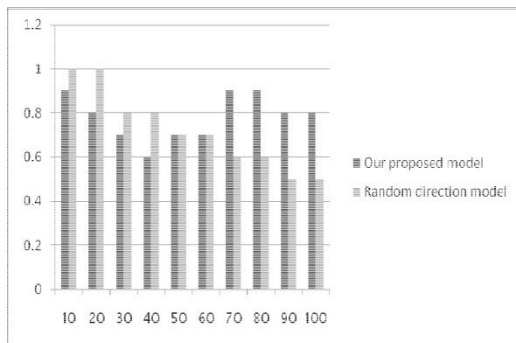


Figure 2: Probability of Link Availability Versus Time.

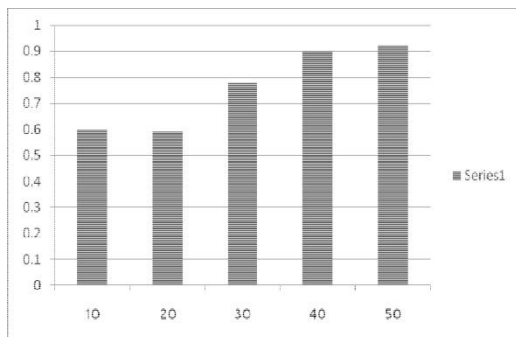


Figure 3: Probability of Link Availability versus Number of Simulated Nodes.

Case 3: Average number of neighbors per node :-

The figure 5 and 6 show the average number of neighbors per node throughout the simulation time. The number of neighbors is being observed for increased level of density. The average number of neighbors per nodes varies smoothly from start to end of the simulation time. At some point in the simulation, the value decreases,

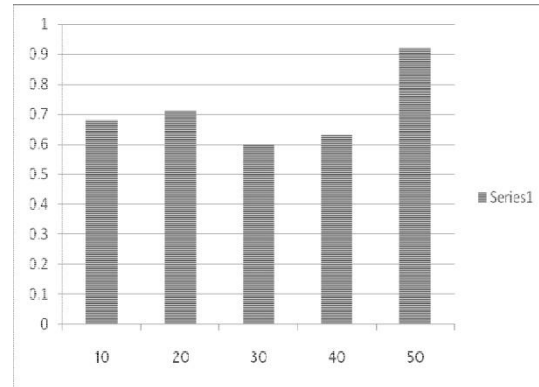


Figure 4: probability of link availability versus number of simulated nodes.

this is due to the intersection of the road. Our model shows that the number of neighbors decreases only at the road intersections, as some nodes change the direction of traveling. With the increase in the density of the nodes, there is increase in the numbers of neighbors per node. This is due to that more number of neighbors per node increase as there are more nodes on the road. Our proposed model gives the better results in terms of average number of neighbors per node. This will give more stability to our model. Our model shows better results in all the experiments, while showing the real situation on the road.

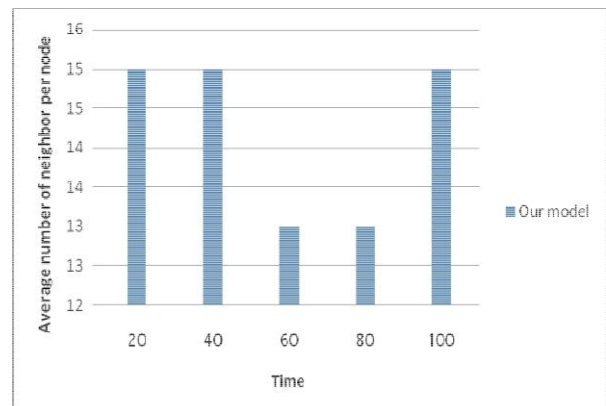


Figure 5: Average numbers of neighbors per node for 50 nodes

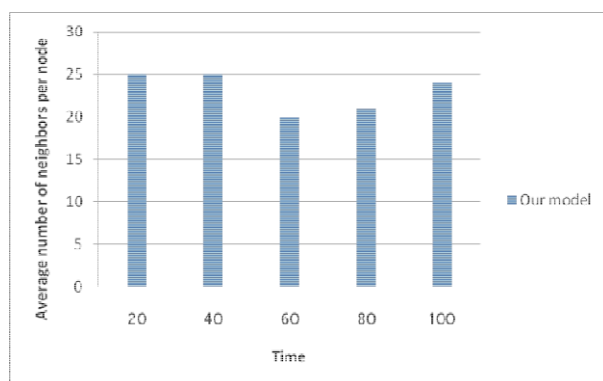


Figure 6: Average numbers of neighbors per node for 75 nodes.

4. Conclusion and Future Work

In this research, we have proposed the new mobility model that covers the city area. In our proposed model we change the speed of nodes after some particular distance, in accordance to neighboring nodes. We assigned probability values to the nodes based on the attraction point, where nodes most likely to move. This would leads to the actual scenarios on the road.

Through simulation we have shown that our model performs better than random direction model in terms probability of link availability. We compared our proposed model with random direction model through the simulation. From the simulation we got the result that shows that probability of link availability decreases at the traffic signal point. This is because of the vehicle either waits or takes turn and change the direction of traveling. Our result demonstrates that probability of link availability in VANETs is more sensitive to the vehicles waiting at intersections and acceleration/deceleration of vehicles. We have tested our model by increasing speed of the nodes. We found that the connectivity among the nodes is very much depends on the speed of the nodes. We also found that performance of mobility model is depends on both the speed and density of the nodes. Thus our model tries to depict the realistic scenarios on the real road.

In future, we will try to extend our model by considering the overtaking parameter into account. We will try to run our model on a two lane of the road.

REFERENCES

- [1] Harri, J.; Filali, F.; Bonnet, C.; "Mobility models for vehicular ad hoc networks: a survey and taxonomy," Communications Surveys & Tutorials, IEEE, vol.11, no.4, pp.19-41, Fourth Quarter 2009
- [2] Azami, M.; Sabaei, M.; Pedram, H., "Adaptive routing protocols for vehicular ad hoc networks," Telecommunications, 2008. IST, 2008. International Symposium on, vol., no., pp.825-830, 27-28 Aug. 2008.
- [3] Yufeng Chen; Zhengao Xiang; Wei Jian; Weirong Jiang; , "An improved AOMDV routing protocol for V2V communication," Intelligent Vehicles Symposium, 2009 IEEE , vol., no., pp.1115-1120, 3-5 June 2009.
- [4] http://en.wikipedia.org/wiki/Mobility_model.
- [5] Mouzna, J.; Uppoor, S.; Boussedjra, M.; Pai, M.M.M., "Density aware routing using road hierarchy for vehicular networks," Service Operations, Logistics and Informatics, 2009. SOLI '09 IEEE/INFORMS International Conference, vol., no., pp.443-448, 22-24 July 2009.
- [6] Vetrisevi, V. and Parthasarathi, R., "Trace based mobility model for ad hoc networks," In Proceedings of the Third IEEE international Conference on Wireless and Mobile Computing, Networking and Communications, WIMOB, IEEE Computer Society, Washington, DC, pp.81-81, 2007.
- [7] Christian Bettstetter, Giovanni Rasta and Paolo Santi, "The Node Distribution of the Random Waypoint Mobility Model for Wireless Ad Hoc Networks", IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 2, NO. 3, JULY-SEPTEMBER 2003.
- [8] Ariyakhajom, Jinthana; Wannawilai, Pattana; Sathitwiriawong, Chanboon; , "A Comparative Study of Random Waypoint and Gauss-Markov Mobility Models in the Performance Evaluation of MANET," Communications and Information Technologies, 2006. ISCIT '06. International Symposium on , vol., no., pp.894-899, Oct. 18 2006-Sept. 20 2006.
- [9] Boundless mobility model, <http://www-public.isudparis.eu/~gauthier/MobilityModel/mobilitymodel.html#Boundless>.
- [10] Zhi Ruxin; Gao Fei; Yang Jie; "Nonuniform Property of Random Direction Mobility Model for MANET," Wireless Communications, Networking and Mobile Computing, 2009. WiCom '09. 5th International Conference on , vol., no., pp.1-4, 24-26 Sept. 2009.
- [11] Ng, J.M.; Yan Zhang; , "Reference region group mobility model for ad hoc networks," Wireless and Optical Communications Networks, 2005.
- [12] F. Kamadi, Z. Mo, K.-C. Lan, "Rapid Generation of Realistic Mobility Models for VANET", Poster Session, 11th Annual International Conference on Mobile Computing and Networking (MobiCom 2005), August 2005.
- [13] SUMO, http://sourceforge.net/apps/mediawiki/sumo/index.php?title=Main_Page.
- [14] D. Choffnes, F. Bustamante, "An Integrated Mobility and Traffic Model for Vehicular Wireless Networks", 2nd ACM Workshop on Vehicular Ad Hoc Networks (VANET 2005), September 2005.
- [15] CANU Project Home Page, <http://canu.informatik.uni-stuttgart.de>.
- [16] <http://www.isi.edu/nsnam/ns/index.html>.
- [17] <http://pcl.cs.ucla.edu/projects/gloimosim/>.
- [18] Gainaru, A.; Dobre, C.; Cristea, V.; "A Realistic Mobility Model Based on Social Networks for the Simulation of VANETs," Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th, vol. no., pp.1-5, 26-29 April 2009.
- [19] M. Treiber, A. Hennecke, D. Helbing, "Congested traffic states in empirical observations and microscopic simulations", Phys. Rev. E62, Issue 2, August 2000.
- [20] Bhandari, Shiddhartha Raj; Lee, Gyu Moun; Crespi, Noel; , "Mobility Model for User's Realistic Behavior in Mobile Ad Hoc Network," Communication Networks and Services Research Conference (CNSR), 2010 Eighth Annual , vol., no., pp.102-107, 11-14 May 2010.

AUTHORS PROFILE



My self Dhananjay Gaikwad, completed Mtech from National Institute of Technology, surat (India) 2011. Since july 2011, I am working as Assistant professor in Parikram college of engineering, kashti. My papers have been published in Springer and IEEE explorer.



My self Mahesh Lagad, pursuing Mtech (MS) from University of Pune (India). I have two year industrial experience, joined the college in july 2011. My paper has been published in International conference.



I, Prashant suryawanshi currenty doing Mtech, from Hyderabad (India). I have total six year Industrial plus academic experience.



I, Vaibhav maske currenty doing Mtech, from Hyderabad (India). I have total four year teaching experience.

Image Classification in Transform Domain

Dr. H. B. Kekre

Professor,
Computer Engineering
Mukesh Patel School of Technology
Management and Engineering,
NMIMS University, Vileparle(w)
Mumbai 400-056, India
hbkekre@yahoo.com.

Dr. Tanuja K. Sarode

Associate Professor,
Computer Engineering,
Thadomal Shahani Engineering
College,
Bandra(W), Mumbai 400-050, India
tanuja_0123@yahoo.com

Jagruiti K. Save

Ph.D. Scholar, MPSTME,
NMIMS University,
Associate Professor,
Fr. C. Rodrigues College of
Engineering, Bandra(W), Mumbai
400-050, India
jagrutik_save@yahoo.com

Abstract— Organizing images into meaningful categories using low level or high level features is an important task in image databases. Although image classification has been studied for many years, it is still a challenging problem within multimedia and computer vision. In this paper the generic image classification approach using different transforms is proposed. The two main steps in image classification are feature extraction and classification algorithm. This paper proposes to generate feature vector from image transform. The paper also investigates the effectiveness of different transforms (Discrete Fourier Transform, Discrete Cosine Transform, Discrete Sine Transform, Hartley and Walsh Transform) in classification task. The size of feature vector also varied to see its impact on the result. Classification is done using nearest neighbor classifier. Euclidean and Manhattan distance is used to calculate the similarity measure. Images from the Wang database are used to carry out the experiments. The experimental results and detailed analysis are presented.

Keywords- Image classification; Image Transform; Discrete Fourier Transform (DFT); Discrete Sine Transform(DST); Discrete Cosine Transform(DCT); Hartley Transform; Walsh Transform; Nearest neighbor Classifier.

I. INTRODUCTION

Though the image classification is usually not a very difficult task for humans, it has been proven to be an extremely complex task for machines. In the existing literatures, most of the frameworks for image classification include two main steps: feature extraction and classification algorithm. In the first step, some discriminative features are extracted to represent the image content such as color [1] [2], shape [3] and texture [4]. There has been a lot of research work done in the area of feature extraction. Saliency map is used to extract features to classify both the query image and database images into attentive and non-attentive classes [5]. The image texture feature is calculated based on gray-level co-occurrence matrix (GLCM) [6]. Color Co-occurrence method in which both the color and texture of an image are taken into account, is used to generate the features [7]. Transforms have been applied to gray scale image to generate feature vector [8]. In classification algorithm step, various multi-class classifiers like k nearest neighbor classifier [9], Support Vector Machine (SVM) [10]

[11], Artificial Neural Network [12] [13], Genetic algorithm [14] are used.

II. IMAGE TRANSFORMS

A. Discrete Fourier Transform (DFT)

The discrete Fourier transform (DFT) is one of the most important transforms that is used in digital signal processing and image processing [15]. Two dimensional discrete Fourier transform for an image $f(x, y)$ of size N by N is given by equation 1.

$$F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \left(\frac{ux}{N} + \frac{vy}{N} \right)} \quad (1)$$

for $0 \leq u, v \leq N-1$

B. Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT), introduced by Ahmed, Natarajan and Rao [16], has been used in many applications of digital signal processing, data compression, information hiding and content based Image Retrieval system(CBIR)[17]. The discrete cosine transform (DCT) is closely related to the discrete Fourier transform. It is a separable linear transformation; that is, the two-dimensional transform is equivalent to a one-dimensional DCT performed along a single dimension followed by a one-dimensional DCT in the other dimension. The two dimensional DCT can be written in terms of pixel values $f(x, y)$ for $x, y = 0, 1, \dots, N-1$ and the frequency-domain transform coefficients $F(u, v)$ as shown in equation 2.

$$F(u, v) = \alpha(u) \alpha(v) \sum_x \sum_y f(x, y) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \quad (2)$$

for $0 \leq u, v \leq N-1$

Where

$$\begin{aligned}\alpha(u) &= 1/\sqrt{N} & \text{for } u = 0 \\ \alpha(u) &= \sqrt{\frac{2}{N}} & \text{for } 1 \leq u \leq N-1 \\ \alpha(v) &= 1/\sqrt{N} & \text{for } v = 0 \\ \alpha(v) &= \sqrt{\frac{2}{N}} & \text{for } 1 \leq v \leq N-1\end{aligned}$$

C. Discrete Sine Transform (DST)

The discrete sine transform was introduced by A. K. Jain in 1974. The two dimensional sine transform is defined by an equation 3.

$$\begin{aligned}F(u,v) &= \frac{2}{N+1} \sum \sum f(x,y) \sin \left[\frac{(x+1)(u+1)\pi}{N+1} \right] \\ &\quad \sin \left[\frac{(y+1)(v+1)\pi}{N+1} \right] \quad (3) \\ &\quad \text{for } 0 \leq u, v \leq N-1\end{aligned}$$

Discrete Sine transform has been widely used in signal and image Processing [18] [19].

D. Discrete Hartley Transform (DHT)

The Hartley transform [20] is an integral transform closely related to the Fourier transform. It has some advantages over the Fourier transform in the analysis of real signals as it avoids the use of complex arithmetic.

A discrete Hartley transform (DHT) is a Fourier-related transform of discrete, periodic data similar to the discrete Fourier transform (DFT), with analogous applications in signal processing and related fields [21]. Its main distinction from the DFT is that it transforms real inputs to real outputs, with no intrinsic involvement of complex numbers. Just as the DFT is the discrete analogue of the continuous Fourier transform, the DHT is the discrete analogue of the continuous Hartley transform. The discrete two dimensional Hartley Transform for image of size $N \times N$ is defined as in equation 4.

$$\begin{aligned}F(u,v) &= \\ \frac{1}{N} \sum \sum f(x,y) \cos \left[\frac{2\pi}{N} (ux + vy) \right] \quad (4) \\ \text{where } \cos \theta &= \cos \theta + \sin \theta\end{aligned}$$

E. Discrete Walsh Transform (DWT))

The Walsh Transform [22] has become quite useful in the applications of image processing [23] [24]. Walsh functions were established as a set of normalized orthogonal functions, analogous to sine and cosine functions, but having uniform values ± 1 throughout their segments. The Walsh transform matrix is defined as a set of N rows, denoted W_j , for $j = 0, 1, \dots, N-1$, which have the following properties:

- W_j takes on the values $+1$ and -1
- $W_j[0] = 1$ for all j
- $W_j \times [W_k]^t = 0$, for $j \neq k$ and $W_j \times [W_k]^t = N$, for $j = k$.
- W_j has exactly j zero crossings, for $j = 0, 1, \dots, N-1$

Each row W_j is even (when j is even) and odd (when j is odd) w.r.t. to its midpoint.

III. ROW MEAN VECTOR

The row mean vector [25] [26] is the set of averages of the intensity values of the respective rows as shown in equation 5.

$$\text{Row mean vector } r = \begin{bmatrix} \text{Avg}(\text{Row } 1) \\ \text{Avg}(\text{Row } 2) \\ : \\ : \\ \text{Avg}(\text{Row } N) \end{bmatrix} \quad (5)$$

IV. PROPOSED ALGORITHM

The image database is divided into a training set and a testing set. The feature vector of each training/testing image is calculated. Given an image to be classified from testing set, a nearest neighbor classifier compares it against the images of a training set, in order to identify the most similar image and consequently the correct class. Euclidean and Manhattan distance is used as similarity measure.

A. Generation of feature vector

1. For each color image $f(x,y)$, generate its three color (R, G, and B) planes $f_R(x,y)$, $f_G(x,y)$ and $f_B(x,y)$ respectively.
2. Apply transform T (DCT, DFT, DST, HARTLEY, WALSH) on the columns of three image planes as given in equation 6 to 8 to get column transformed images.

$$[T] \times [f_R(x,y)] = F_R(x,v) \quad (6)$$

$$[T] \times [f_G(x,y)] = F_G(x,v) \quad (7)$$

$$[T] \times [f_B(x,y)] = F_B(x,v) \quad (8)$$

3. Calculate row mean vector of each column transformed image.
4. Make a feature vector of size 75 by fusing the row mean vectors of R, G, and B plane. Take first 25 values from R plane followed by first 25 values from G plane followed by first 25 values from B plane.

5. Do the above process for training images to generate the feature database.

The different values of feature vector size like 150 (50R + 50G + 50B), 225 (75R + 75G + 75B), 300 (100R + 100G + 100B), 450 (150R + 150G + 150B), and 768 (256R + 256G + 256B) are also considered to generate feature vectors.

B. Classification

1. In this phase, for given testing images, their feature vectors are generated.
2. Euclidean distance and Manhattan distance is calculated between each testing image feature vector and each training image feature vector.
3. Minimum distance indicates the most similar training image for that testing image. Then the given testing image is assigned to the corresponding class.

We have also considered another training set where each feature vector is the average of feature vectors of all training images of a particular class.

V. RESULTS

The implementation of the proposed technique is done in MATLAB 7.0 using a computer with Intel Core 2 Duo Processor T8100 (2.1GHz) and 2 GB RAM. The proposed technique is tested on the Wang image database. This database was created by the group of professor Wang from the Pennsylvania State University [27]. The experiment is carried on 8 classes of Wang database. For testing, 30 images for each class were used and for training, 5 images of each class were used. Thus total testing images were 240 and total training images were 40. Training set contains 40 feature vectors. The proposed method is also implemented using another training set that contain 8 feature vectors where each feature vector is the average of feature vectors of all training images of same class. Fig. 1 shows the sample database of training images and Fig. 2 shows the sample database of testing images.



Figure 1. Sample database of training images

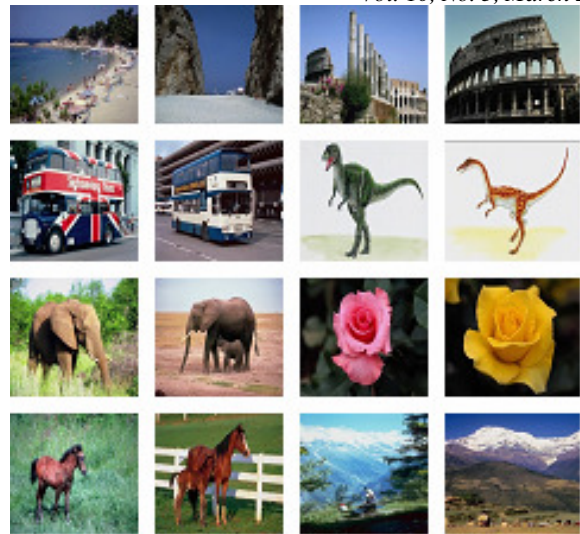


Figure 2. Sample database of testing images

Each image is resized to 256 x 256. Table I and Table II shows the number of correctly classified total images (out of 240) for different transforms over different vector sizes for two different training sets. The correctness of classification is visually checked.

With average training set Walsh transform gives better performance compared to other transforms with Manhattan as similarity measure. If Euclidean distance is used for calculation then feature vector size of 768 gives the marginally better performance in all transforms. Considering the results as shown in Table 1, best results are obtained for Manhattan distance as similarity measure. DST Walsh and DFT gave better performance in that order.

Now considering individual class classification performance using these two similarity measures is shown in Table III to Table VI. For this purpose the vector size is selected based on the performance. For Euclidean distance criterion, the number of correctly classified training images in each class for different transforms over two training sets is shown in table III and table IV with feature vector size 768. If a Manhattan distance criterion is used, then there is a variation in the performance of the transforms for different feature vector sizes. In most cases vector size 225 gives better performance. So using this vector size, the number of correctly classified images in each class for different transforms over two training sets is shown in table V and table VI.

TABLE I. NUMBER OF CORRECTLY CLASSIFIED IMAGES (OUT OF 240) FOR DFT, DCT, DST, HARTLEY AND WALSH OVER DIFFERENT FEATURE VECTOR SIZES USING EUCLIDEAN AND MANHATTAN DISTANCE., TRAINING SET: FEATURE VECTORS OF 5 IMAGES FROM EACH CLASS

Transform	Distance E-Euclidean M-Manhattan	Feature vector size					
		75	150	225	300	450	768
DFT	E	155	159	159	159	160	167
	M	166	163	169	169	164	163
DCT	E	151	156	159	162	162	163
	M	163	167	169	170	164	163
DST	E	159	160	160	160	161	160
	M	164	173	176	174	168	161
HARTLEY	E	148	150	151	151	152	158
	M	154	162	165	167	161	161
WALSH	E	149	152	155	156	160	161
	M	160	162	166	170	171	170

TABLE II. NUMBER OF CORRECTLY CLASSIFIED IMAGES (OUT OF 240) FOR DFT, DCT, DST, HARTLEY AND WALSH OVER DIFFERENT FEATURE VECTOR SIZES USING EUCLIDEAN AND MANHATTAN DISTANCE., TRAINING SET: AVERAGE OF FEATURE VECTORS OF 5 IMAGES FROM EACH CLASS

Transform	Distance E-Euclidean M-Manhattan	Feature vector size					
		75	150	225	300	450	768
DFT	E	155	160	162	162	161	166
	M	175	173	171	169	164	156
DCT	E	156	158	157	159	160	160
	M	171	172	169	168	163	156
DST	E	161	160	160	159	161	161
	M	161	162	168	169	169	164
HARTLEY	E	159	162	161	162	163	167
	M	169	168	172	171	168	164
WALSH	E	155	157	158	158	158	159
	M	179	175	173	169	169	159

TABLE III. TOTAL CLASSIFIED IMAGES (OUT OF 30 IMAGES) IN EACH CLASS FOR DIFFERENT TRANSFORMS, VECTOR SIZE: 768, DISTANCE CRITERIA: EUCLIDEAN DISTANCE, TRAINING: FEATURE VECTORS OF 5 IMAGES FROM EACH CLASS

Classes	DFT	DCT	DST	HARTLEY	WALSH
Beach	15	14	11	14	11
Monument	10	13	7	9	8
Bus	24	21	27	22	25
Dinosaur	30	30	30	30	30
Elephant	24	23	23	24	24
Flower	27	25	26	27	25
Horse	26	28	26	25	28
Snow Mountain	11	9	10	7	10

TABLE IV. TOTAL CLASSIFIED IMAGES (OUT OF 30 IMAGES) IN EACH CLASS FOR DIFFERENT TRANSFORMS, VECTOR SIZE: 768, DISTANCE CRITERIA: EUCLIDEAN DISTANCE, TRAINING: AVERAGE OF FEATURE VECTORS OF 5 IMAGES FROM EACH CLASS

Classes	DFT	DCT	DST	HARTLEY	WALSH
Beach	20	18	14	19	17
Monument	3	4	9	6	5
Bus	23	24	25	23	24
Dinosaur	30	30	30	30	30
Elephant	25	22	24	25	24
Flower	30	30	29	30	30
Horse	16	17	17	16	16
Snow Mountain	19	15	13	18	13

TABLE V. TOTAL CLASSIFIED IMAGES (OUT OF 30 IMAGES) IN EACH CLASS FOR DIFFERENT TRANSFORMS, VECTOR SIZE: 225, DISTANCE CRITERIA: MANHATTAN DISTANCE, TRAINING: FEATURE VECTORS OF 5 IMAGES FROM EACH CLASS

Classes	DFT	DCT	DST	HARTLEY	WALSH
Beach	23	21	19	24	23
Monument	9	11	9	11	8
Bus	25	20	27	22	24
Dinosaur	30	30	30	30	30
Elephant	22	23	20	22	21
Flower	30	28	30	30	25
Horse	22	23	24	19	25
Snow Mountain	8	13	17	7	10

TABLE VI. TOTAL CLASSIFIED IMAGES (OUT OF 30 IMAGES) IN EACH CLASS FOR DIFFERENT TRANSFORMS, VECTOR SIZE: 225, DISTANCE CRITERIA: MANHATTAN DISTANCE, TRAINING SET: AVERAGE OF FEATURE VECTORS OF 5 IMAGES FROM EACH CLASS

Classes	DFT	DCT	DST	HARTLEY	WALSH
Beach	24	23	16	24	26
Monument	9	9	7	11	6
Bus	24	25	26	25	28
Dinosaur	30	30	30	30	30
Elephant	21	18	21	21	19
Flower	30	30	30	30	30
Horse	20	22	22	20	22
Snow Mountain	13	12	16	11	12

The comparisons of performances of different transforms are shown in Fig. 3 to Fig. 6.

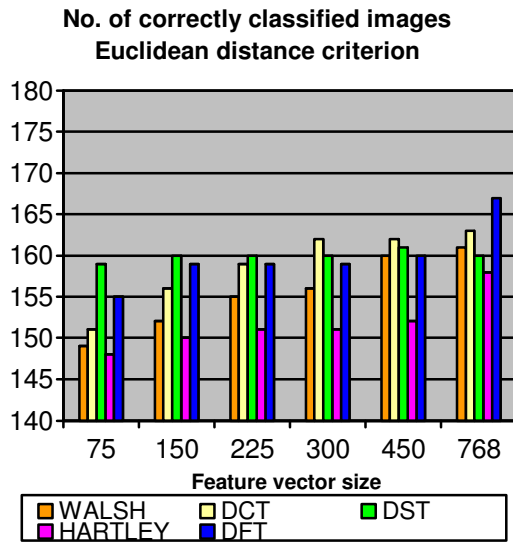


Figure 3. Performance of different transform (training set: Feature vectors of 5 images from each class)

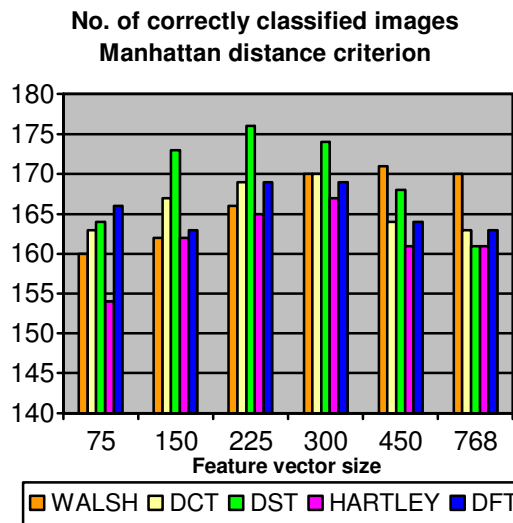


Figure 4. Performance of different transform (training set: Feature vectors of 5 images from each class)

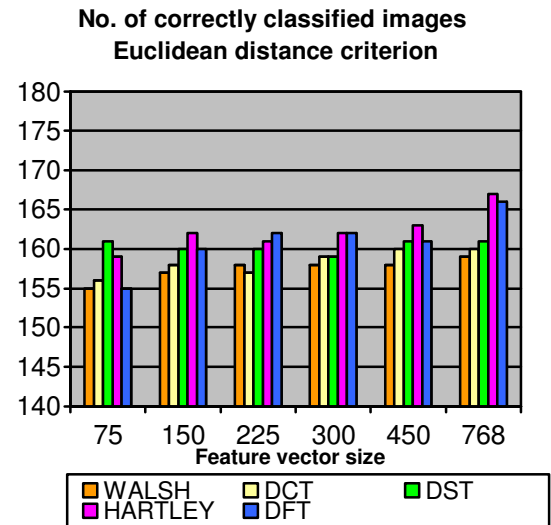


Figure 5. Performance of different transform (training set: Average of feature vectors of 5 images from each class)

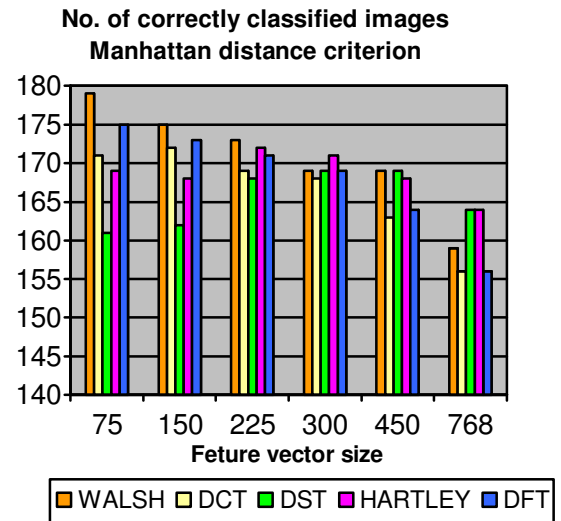


Figure 6. Performance of different transform (training set: Average of feature vectors of 5 images from each class)

VI. CONCLUSIONS

This paper proposes to prepare the feature vector from an image column transform and use it for image classification. This gives considerable saving of computational time as compared to full transform. The paper investigates the performance of different transforms. The performance is tested thoroughly using different criteria like distance measure (Euclidean distance, Manhattan distance); size of feature vector (75, 150, 225, 300, 450 and 768) and training sets (feature vectors, average of feature vectors). Conclusion

from the results of individual class classification is given in Table VII.

TABLE VII. BEST 3 CLASS PERFORMANCES FOR DIFFERENT CRITERIA

Training Set	Similarity Measure	Best 3 performer classes
Feature vectors of 5 images from each class	Euclidean	Dinosaur (100%) Horse (88.66%) Flower (86.66%)
	Manhattan	Dinosaur (100%) Flower (95.33%) Bus (78.66%)
Average of feature vectors of 5 images from each class	Euclidean	Dinosaur (100%) Flower (99.33%) Elephant (80%)
	Manhattan	Dinosaur (100%) Flower (100%) Bus (85.33%)

Results also show that the training set containing average of feature vectors, gives better results and since they are less in numbers, the computation is fast. It is also seen that Manhattan distance gives high performance for small feature vector size when compared with Euclidean distance criterion.

REFERENCES

- [1] M. J. Swain and D. H. Ballard, "Color indexing," International Journal of Computer Vision, vol.7, no.1, pp.11-32, 1991.
- [2] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," Pattern recognition, vol.29, no.8, pp.1233-1244, 1996
- [3] F. Mokhtarian and S. Abbasi, "Shape similarity retrieval under affinetransforms," Pattern Recognition, 2002, vol. 35, pp.31-41.
- [4] B.S.Manjunath and W.Y.Ma, "Texture feature for browsing and retrieval of image data," IEEE Pattern Analysis and Machine Intelligence, no. 18, vol. 8, pp. 837- 842, 1996.
- [5] Z. Liang, H. Fu, Z. Chi, and D. Feng, "Image Pre-Classification Based on Saliency Map for Image Retrieval," Proc. of the IEEE International Conference on Information, Communications and Signal Processing, pp. 1-5, Dec 2009.
- [6] F. Siraj, M. Salahuddin, and S. Yusof, "Digital Image Classification for Malaysian Blooming Flower," the IEEE Second International Conference on Computational Intelligence, Modelling and Simulation, (CIMSIm), pp. 33-38,Bali, Sept 2010.
- [7] D. Bashish, M. Braik, and S. Bani-Ahmad, "A Framework for Detection and classification of Plant Leaf and Stem Diseases," Proc. of the IEEE International Conference on signal and image processing (ICSIP), pp. 113-118, Chennai Dec 2010.
- [8] H.B. Kekre, T. K. Sarode, M. S. Ugale, "Performance Comparison of Image Classifier Using DCT, Walsh, Haar and Kekre's Transform," International Journal of Computer Science and Information Security,(IJCSIS), Vd9, No. 7, 2011
- [9] M. Szummer and R. W. Picard, "Indoor-Outdoor Classification," IEEE International workshop Content based Access of Image and Video Databases, in conjunction with ICCV'98, pp. 384-390, Jan 2009.
- [10] O. Chapelle, P. Haffner, and V. Vapnik, "Support vector machines for histogram- based image classification," IEEE Transactions on Neural Networks, vol. 10, pp. 1055-1064, 1999.
- [11] S. Agrawal, N. Verma, P. Tamrakar, and P. Sircar, "Content Based Color Image Classification using SVM," in Proc. of IEEE International Conference on Information Technology: New Generations (ITNG), pp. 1090 – 1094, Las Vegas, April 2011.
- [12] M. Lotfi1, A. Solimani, A. Dargazany, H. Afzal, and M. Bandarabadi, "Combining wavelet transforms and neural networks for image

classification," the IEEE Symposium on System Theory, SSST, pp.44-48, Aug 2009.

- [13] S. Sadek, A. Hamadi, B. Michaelis, and U. Sayed, "Robust Image Classification Using Multi-level Neural Networks," Proc. of the IEEE International Conference on Intelligent Computing and Intelligent Systems, Vol.: 4, pp. 180 – 183, Shanghai Dec 2009.
- [14] J. Z. Wang, J. Li and G. Wiederhold, "SIMPLiCity: semantic sensitive integrated matching for picture libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, vol.23, no.9, pp.947-963.
- [15] E. O. Brigham, R. E. Morrow, "The Fast Fourier Transform," Spectrum, IEEE, Dec. 1967, Vol. 4, Issue 12, pp. 63-70.
- [16] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete Cosine Transform," IEEE Transactions, Computers, 90-93, Jan 1974.
- [17] H. B.Kekre, T. K. Sarode, S. D. Thepade, "Color-Texture Feature based Image Retrieval using DCT applied on Kekre's Median Codebook", International Journal on Imaging (IJI), Volume 2, Number A09, Autumn 2009, pp. 55-65. Available online at www.ceser.res.in/iji.html (ISSN: 0974-0627) .
- [18] S. A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," IEEE Transactions on Signal Processing, Vol. 42, Issue 5, pp. 1038-1051, 1994.
- [19] H. B.Kekre and D. Mishra, "Feature Extraction of Color Images using Sectorization of Discrete Sine Transform," IJCA Proceedings on International Conference and workshop on Emerging Trends in Technology (ICWET), Vol. 4, pp.:27-32, 2011.
- [20] Hartley, R. V. L., "A More Symmetrical Fourier Analysis Applied to Transmission Problems," Proceedings IRE 30, pp.144–150, Mar-1942.
- [21] R. P. Millane, "Analytical properties of the Hartley Transform and its Implications", Proceedings of the IEEE, Mar. 1994, Vol. 82, Issue 3, pp. 413-428.
- [22] J. L.Walsh, "A Closed Set of Orthogonal Functions," American Journal of Mathematics, vol. 45, pp. 5-24, 1923 .
- [23] H. B.Kekre and D. Mishra, "Density Distribution and Sector Mean with Zero-Sal and Highest-Cal Components in Walsh transform Sectors as Feature Vectors for Image Retrieval," International Journal of Computer Science and Information Security (IJCSIS), vol.8, No. 4, 2010, ISSN 1947-5500.
- [24] H. B.Kekre, Vinayak Bharadi, "Walsh Coefficients of the Horizontal & Vertical Pixel Distribution of Signature Template", In Proc. of Int. Conference ICIP-07, Bangalore University, Bangalore. 10-12 Aug 2007.
- [25] H. B.Kekre, Sudeep D. Thepade, Akshay Maloo "Performance Comparison for Face Recognition using PCA, DCT &WalshTransform of Row Mean and Column Mean", ICGST International Journal on Graphics, Vision and Image Processing (GVIP), Volume 10, Issue II, pp.9-18, June 2010.
- [26] H.B.Kekre, Tanuja Sarode, Sudeep D. Thepade, "DCT Applied to Row Mean and Column Vectors in Fingerprint Identification", In Proceedings of Int. Conf. on Computer Networks and Security (ICCNS), 27-28 Sept. 2008, VIT, Pune.
- [27] Wang, J. Z., Li, J., Wiederhold, G.: SIMPLiCity: Semantics-sensitive Integrated Matching for Picture Libraries, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 23, no.9, pp. 947-963, (2001).

AUTHORS PROFILE



Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engineering, from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay.

For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering. College, Mumbai.

Now he is Senior Professor at MPSTME, SVKM's NMIMS University. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E./ B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 450 papers in National /International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE Recently twelve students working under his guidance have received best paper awards and six research scholars have been conferred Ph. D. Degree by NMIMS University. Currently 7 research scholars are pursuing Ph.D. program under his guidance.



Tanuja K. Sarode has Received Bsc. (Mathematics) from Mumbai University in 1996, Bsc.Tech.(Computer Technology) from Mumbai University in 1999, M.E. (Computer Engineering) from Mumbai University in 2004, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 10 years of experience in teaching. Currently working as Associate Professor in

Dept. of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is life member of IETE, ISTE, member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT), Singapore. Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has more than 100 papers in National /International Conferences/journal to her credit.



Jagruti K. Save has received B.E. (Computer Engg.) from Mumbai University in 1996, M.E. (Computer Engineering) from Mumbai University in 2004, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 10 years of experience in teaching. Currently working as Associate Professor in Dept. of Computer Engineering at Fr. Conceicao Rodrigues College of Engg., Bandra, Mumbai. Her areas of interest are Image Processing, Neural Networks, Fuzzy systems, Data base management and Computer Vision. She has 6 papers in National /International Conferences/journal to her credit.

Analysis of Stock Marketing with SOAP service using Python

P.Asha¹

¹Research Scholar, Computer Science and Engineering Department, Sathyabama University, Chennai, Tamilnadu, India.
ashapandian225@gmail.com

Dr.T.Jebarajan²

²Principal, Kings College of Engineering, Chennai, Tamilnadu, India.
drtjebarajan@gmail.com

Kathiresan³

³Technical Lead Consultant, Motorola Solutions, Bangalore, Karnataka, India
kathir.it@gmail.com

Abstract - SOAP is a simple XML-based protocol specification to let applications exchange information over HTTP. SOAP describes envelope and message formats, and has a basic request/response handshake protocol. A SOAP message could be sent to a web-service-enabled web site such as a real-estate price database, with the parameters needed for a search. The site would then return an XML formatted document with the resulting data, e.g., prices, location, features. With the data being returned in a standardized machine-parseable format, it can then be integrated directly into a third party web site or application. SOAPpy provides tools for building SOAP clients and servers. SOAPpy is very simple to use and that fully supports dynamic interaction between clients and servers.

Keywords— Service Oriented Architecture, SOAP, Web Service, SOAPpy, XML.

I. INTRODUCTION

An **architectural style** is a coordinated set of architectural constraints that restricts the roles/features of architectural elements and the allowed relationships among those elements within any architecture that conforms to that style.

There are different Software Architecture Styles:

- Data Oriented Architecture
- Hierarchical Architecture
- Call and Return Architecture
- Interaction Process Architecture
- Service Oriented Architecture
- Space based Architecture

II. SERVICE ORIENTED ARCHITECTURE

A service is a function that is well-defined, self-contained, and does not depend on the context or state of other services. A service-oriented architecture is a collection of services. These services communicate with each other which involve either simple data passing or it could involve two or more

services coordinating some activity. Some means of connecting services to each other is needed.

A web service is the connection technology of service-oriented architectures.

- “A **Web Service (WS)** is a software system designed to support *interoperable machine-to-machine* interaction over a network.”
- Usually a WS provides the **API (Application Programming Interface)**.

Web services essentially use **XML** to create a robust connection.

A. SOA Interaction Pattern

- A service provider creates a service for interaction and exposes the service's description for the consumers with the necessary message format and transport bindings.

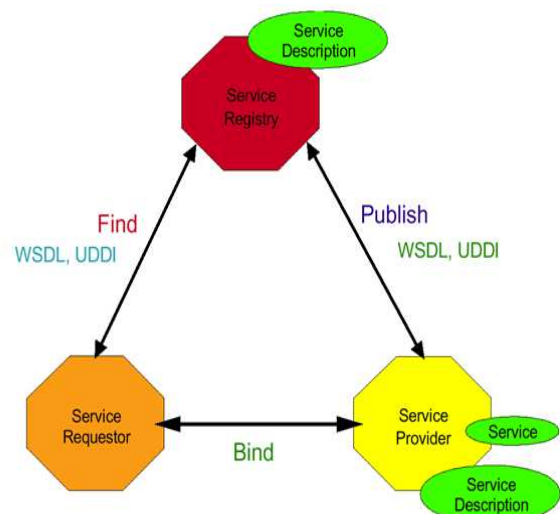


Figure 1. Service Oriented Architecture model implemented by XML Web Services

- The service provider may decide to register this service and its description with a registry of choice.
- The service consumer can discover a service from a registry or directly from the service provider and can start sending messages in a well-defined XML format that both the consumer and service can consume.

B. SOA Realization (Two ways)

SOA can be realized by the following ways:

- XML - SOAP Based Web Services
- ReSTful Web services

Earlier, the message exchange between a consumer and producer was by using a common, well understandable, and interoperable data model, HTML/XHTML.



Figure 2. WebSites (1992)

Later on, when the interaction pattern becomes complex, such as business-to-business, the above Web architecture model needs more polished message exchange patterns to adapt to any user agent and/or applications of choice. The Web service architecture extends the above interaction pattern further by adding the power and expressiveness of XML.

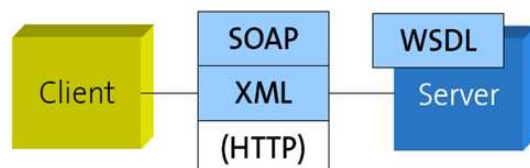


Figure 3. WS - * Web Services (later)

The message exchange centered on XML and the interaction pattern have evolved to an any-to-any scenario. This flexible interaction pattern using XML messages as the core data format increases the value of SOAs. This produces the interaction request-response patterns, asynchronously, for better interoperability between consumers and any of its producers. This is a very loosely coupled system.

III. SIMPLE OBJECT ACCESS PROTOCOL

SOAP (Simple Object Access Protocol) is required for application development to allow Internet communication between programs. SOAP provides a way to communicate between applications running on different operating systems, with different technologies and programming languages.

A. SOAP Processing Model and Message Format

SOAP sender - A SOAP node that transmits a SOAP message.

SOAP receiver - A SOAP node that accepts a SOAP message.

SOAP message path - The set of SOAP nodes through which a single SOAP message passes.

Initial SOAP sender (Originator) - The SOAP sender that originates a SOAP message at the starting point of a SOAP message path.

SOAP intermediary - A SOAP intermediary is both a SOAP receiver and a SOAP sender and is targetable from within a SOAP message. It processes the SOAP header blocks targeted at it and acts to forward a SOAP message towards an ultimate SOAP receiver.

Ultimate SOAP receiver - The SOAP receiver that is a final destination of a SOAP message. It is responsible for processing the contents of the SOAP body and any SOAP header blocks targeted at it.

B. SOAPpy

SOAPpy is a SOAP-1.1 library for Python which uses WSDL and SDL documents to discover SOAP-based service APIs. It also includes an XML Schema parser which can parse a subset of the XML Schema standard.

Features of SOAPpy

- Automatic stateful SOAP server support
- SOAP 1.0
- WSDL client/server support
- SSL clients/servers (based on OpenSSL)
- General SOAP Parser/Builder based on sax.xml
- SOAP for RPC client/server code

IV. CASE STUDY

Here, we try to develop a Stock Market SOAP web service to get today's sensx value from SOAP server.

Steps:

1. Develop a Server function [stkmarket](#) in Python to calculate today's sensx value:

- a. Store yesterday's total market capitalization of 30 companies in `yst_totmktcap` and yesterday's sensx value in `yst_sensex`.
 - b. Get current market capitalization of 30 companies in a dictionary.
 - c. Add current market capitalization of 30 companies to find current total market capitalization (`totmktcap`).
 - d. Find today's sensx value using the formula:
$$\text{sensex} = \text{yst_sensex} * \text{totmktcap} / \text{yst_totmktcap}$$
 - e. Find the difference of sensx values `yst_sensex` and `sensex` to show Market is UP or DOWN.
 - f. Return the result string with today's sensx value and market status.
2. Create an object for SOAPServer to listen on "localhost" and port number 8080.
 3. Register the Server function `stkmarket` in SOAPServer.
 4. Start the Stock Market Web Service by listening with the SOAPServer object.
 5. Develop a Stock Market Client in Python to utilize the Stock Market Web Service:
 - a. Create an object for SOAPProxy with `http://localhost:8080/`
 - b. Enable debug messages to print the SOAP request and response messages in the console.
 - c. Connect to the Stock Market Server and utilize the `stkmarket` webservice.
 6. Display today's sensx value and market status in the Client.

1) Stock Market SOAP Server - soapserver.py

```
# soapserver.py
import SOAPpy
def stkmarket():
    yst_totmktcap = 2846906.42
    yst_sensex = 16990.18

    #market capitalization for 30
    companies
    mktcap = {
        'mnf':18189.9, 'dlf':34070.83,
        'itc':154745.53, 'hdfc':97137.93,
        'jaipra':93983.21,
        'bajajaut':41301.80, 'maruti':34895.
        76, 'ongc':244857.84,
        'bhel':84760.39,
        'hdfcbank':108904.34,
        'sbi':141640.87, 'hul':70719.84,
        'heromotoco':37054.34,
        'bharatiatel':154597.04,
        'ntpc':139636.87,
        'jindalstell':48467.84,
        'icicibank':108414.38,
        'lnt':98332.13,
        'sterliteind':45364.79, 'tatapower':
        28010.89, 'ril':250612.63,
```

```
'hindalco':28296.31,
'sunpharma':50174.82,
'coalindia':240400.66,
'wipro':85298.49, 'cipla':55534.05,
'infy':136488.33, 'tcs':189310.17,
'tatamotors':50490.38,
'tatasteel':46622.40
}
totmktcap = sum([i for i in
mktcap.values()])
sensex = yst_sensex * totmktcap /
yst_totmktcap

str = 'UP' if ((sensex -
yst_sensex) > 0) else 'DOWN'

res = "%s %d (%d %s)" % ('Today
sensex value is', sensx,
sensex - yst_sensex, str)
return res

server=
SOAPpy.SOAPServer(("localhost",
8080))
server.registerFunction(stkmarket)
print 'Stock Market SOAP Server is
listening...'
server.serve_forever()
```

2) Stock Market SOAP Client - soapclient.py

```
# soapclient.py
import SOAPpy
server =
SOAPpy.SOAPProxy("http://localhost:
8080/")
server.config.debug = 1
print 'Stock Market SOAP Client is
connecting...'
print server.stkmarket()
```

V. RESULTS

OUTPUT

```
In build.
Outgoing HTTP headers
*****
POST / HTTP/1.0
Host: localhost:8080
User-agent: SOAPpy 0.12.4
(http://pywebsvcs.sf.net)
Content-type:text/xml;charset=UTF-8
Content-length: 348
SOAPAction: "stkmarket"
*****
OutgoingSOAP
*****
<?xml version="1.0" encoding="UTF-
8"?>
<SOAP-ENV:Envelope SOAP-
ENV:encodingStyle="http://schemas.x
mlsoap.org/soap/encoding/"
```

```

xmlns:SOAP-
ENC=http://schemas.xmlsoap.org/soap/encoding/
ENV=http://schemas.xmlsoap.org/soap/envelope/
<SOAP-ENV:Body>
<stkmarket SOAP-ENC:root="1">
</stkmarket>
</SOAP-ENV:Body>
</SOAP-ENV:Envelope>
*****
code= 200    msg= OK    headers=
Server:<a
href=http://pywebsvcs.sf.net>
SOAPpy 0.12.4</a> (Python 2.7.1)
Date: Wed, 10 Aug 2011 11:09:10 GMT
Content-type: text/xml; charset=UTF-8
Content-length: 543
content-type=text/xml;
charset=UTF-8
data=      <?xml      version="1.0"
encoding="UTF-8"?>
<SOAP-ENV:EnvelopeSOAP-
ENV:encodingStyle=http://schemas.xmlsoap.org/soap/encoding/
xmlns:SOAP-
ENC=http://schemas.xmlsoap.org/soap/encoding/
xmlns:xsi=http://www.w3.org/1999/XMLSchema-instance
xmlns:SOAP-
ENV=http://schemas.xmlsoap.org/soap/envelope/

```

```

xmlns:xsd=http://www.w3.org/1999/XMLSchema
<SOAP-ENV:Body>
<stkmarketResponseSOAP-
ENC:root="1">
<Result xsi:type="xsd:string">
Today sensex value is 17416
(426 UP)</Result>
</stkmarketResponse>
</SOAP-ENV:Body>
</SOAP-ENV:Envelope>

```

```

Incoming      HTTP      headers
*****
HTTP/1.? 200 OK
Server:<a
href=http://pywebsvcs.sf.net>SOAP
py 0.12.4</a> (Python 2.7.1)
Date: Wed, 10 Aug 2011 11:09:10 GMT
Content-type: text/xml; charset=UTF-8
Content-length: 543
*****
Incoming      SOAP
*****
<?xmlversion="1.0"encoding="UTF-
8"?>
<SOAP-ENV:Envelope
SOAP-ENV:encodingStyle=
http://schemas.xmlsoap.org/soap/encoding/
xmlns:SOAP-
ENC=http://schemas.xmlsoap.org/soap/encoding/

```

```

xmlns:xsi=http://www.w3.org/1999/XMLSchema-instance

```

```

xmlns:SOAP-
ENV=http://schemas.xmlsoap.org/soap/envelope/

```

```

xmlns:xsd=http://www.w3.org/1999/XMLSchema

```

```

<SOAP-ENV:Body>
<stkmarketResponseSOAP-
ENC:root="1">
<Result xsi:type="xsd:string">Today
sensex value is 17416 (426
UP)</Result>
</stkmarketResponse>
</SOAP-ENV:Body>
</SOAP-ENV:Envelope>
*****

```

Today sensex value is 17416 (426 Up)

Hence simulation of SOAP messages for stock market sensex calculation is made.

Client asks server about today's sensex value. In server, we have a dictionary of key-value pairs (company and their market capitalization value). Based on these values the sensex is calculated. The calculated sensex value is returned back to the client. All the communication happens as a SOAP message. Though we are doing RPC using python, the python SOAP library (SOAPpy) makes this RPC with SOAP messages.

At server side, one python function which is registered in SOAPpy server - these provides as the web service description and publish. At client side, the SOAPpy proxy client invokes the server function (as a RPC call) - this simulate the discovery of web service and the message in between the entities are send as a SOAP envelopes with header and body blocks.

VI. CONCLUSION

A Stock Market SOAP web service has been developed to get today's sensex value from SOAP server and the results were produced. Simulation of Web service (without WSDL) is done and not a form of registration in UDDI is made. Moreover there is no real time service description, publish and discovery, everything is taken care of by SOAPpy.

REFERENCES

- [1] D. Braga, A. Campi, S. Ceri, M. Klemettinen, PL. Lanzi, "Mining Association Rules from XML Data", in Proceedings of DEXA 2002 (DaWaK), LNCS 2454, Aix-en-Provence, France, Sep. 2002, pp. 21-30.

- [2] Potts, M. *Find Bind and Execute: Requirements for Web Service Lookup and Discovery*.
www.talkingblocks.com/resources.htm#,
accessed January 2003.
- [3] Mohammed J. Zaki, Charu C. Aggarwa, "XRules: An Effective Structural Classifier for XML Data", Rensselaer Polytechnic Institute.
- [4] <http://www.rediff.com/money/2000/apr/26sspe.htm>
- [5] <http://www.informit.com/articles/article.aspx?p=336265>
- [6] Locad.com: Web Services Tutorial with Python.
<<http://www.lokad.com/web-services-time-seriesforecasting-tutorial-python.ashx>>, January 2, 2008.
- [7] World Wide Web Consortium (W3C): <http://www.w3.org>
- [8] Joukl, Holger: Interoperable WSDL/SOAP web services introduction: Python ZSI, Excel XP, gSOAP/C++ & ApplixSS).<http://pywebsvcs.sourceforge.net/holger.pdf>>. July 22, 2005.
- [9] Boverhof, Joshua; Moad, Charles: ZSI: The Zolera Soap Infrastructure User's Guide.
<<http://downloads.sourceforge.net/pywebsvcs/ZSI-2.1-a1.tar.gz>>, Release 2.1.0, November 01, 2007.
- [10] WebServices.org:<http://www.webservices.org>

Accurate Face Recognition Using PCA and LDA

Sukhvinder Singh*
Mtech CSE (4th sem)
Sri Sai College Of Engg. & Tech.,
Pathankot
sukhaish@gmail.com

Meenakshi Sharma
HOD CSE
Sri Sai College Of Engg. & Tech.,
Pathankot
mss.s.c.e.t@gmail.com

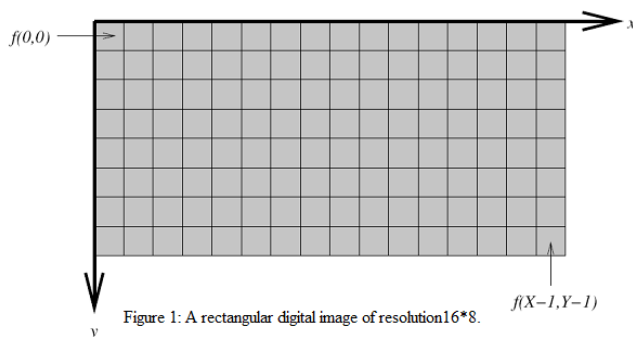
Dr. N Suresh Rao
HOD CSE
Sri Sai College Of Engg. & Tech.,
Jammu University

Abstract: Face recognition from images is a sub-area of the general object recognition problem. It is of particular interest in a wide variety of applications. Here, the face recognition is based on the new proposed modified PCA algorithm by using some components of the LDA algorithm of the face recognition. The proposed algorithm is based on the measure of the principal components of the faces and also to find the shortest distance between them. The experimental results demonstrate that this arithmetic can improve the face recognition rate. . Experimental results on ORL face database show that the method has higher correct recognition rate and higher recognition speeds than traditional PCA algorithm.

Keywords: Face recognition, PCA, LDA.

I. INTRODUCTION

A digital image is a discrete two-dimensional function $f(x,y)$ which has been quantized over its domain and range . Without loss of generality, it will be assumed that the image is rectangular, consisting of x rows and y columns.[13] The resolution of such an image is written as $x*y$. By convention, $f(0,0)$ is taken to be the top left corner of the image, and $f(x-1,y-1)$ the bottom right corner. This is summarized in Figure 1.



Each distinct coordinate in an image is called a pixel, which is short for picture element. The nature of the output of $f(x,y)$ for each pixel is dependent on the type of image. Most images are the result of measuring a specific physical phenomenon, such as light, heat, distance, or energy. The measurement could take any numerical form. A greyscale image measures light intensity only. Each pixel is a scalar proportional to the brightness. The minimum

brightness is called black, and the maximum brightness is called white. A typical example is given in Figure 2.[15] A colour image measures the intensity and chrominance of light. Each colour pixel is a vector of colour components. Common colour spaces are RGB (red, green and blue), HSV (hue, saturation, value), and CMYK (cyan, magenta, yellow, black), which is used in the printing industry. Pixels in a range image measure the depth of distance to an object in the scene[30]. Range data is commonly used in machine vision applications.



Figure 2: A typical greyscale image of resolution 512*512.

For storage purposes, pixel values need to be quantized. The brightness in greyscale images is usually quantized to levels, so $f(x,y)$ belongs to $\{0, 1, \dots, z-1\}$. If z has the form 2^L the image is referred to as having L bits per pixel. Many common greyscale images use 8 bits per pixel giving 256 distinct grey levels. This is a rough bound on the number of different intensities the human visual system is able to discern. For the same reasons, each component in a colour pixel is usually stored using 8 bits [17].

Medical scans often use 12-16 bits per pixel, because their accuracy could be critically important. Those images to be processed predominantly by machine may often use higher values to avoid loss of accuracy throughout processing. Images not encoding visible light intensity, such as range data, may also require a larger value of z to store sufficient distance information.

There are many other types of pixels. Some measure bands of the electromagnetic spectrum such as infra-red or radio, or heat, in the case of thermal images. Volume images are actually three dimensional images, with each pixel being called a voxel. In some cases, volume images may be treated as adjacent two-dimensional image slices. [43] Although this thesis deals with grayscale images, it is often straightforward to extend the methods to function with different types of images.

II. Recognition

Face recognition from images is a sub-area of the general object recognition problem. It is of particular interest in a wide variety of applications. Applications in law enforcement for mugshot identification, verification for personal identification such as driver's licenses and credit cards, gateways to limited access areas, surveillance of crowd behavior are all potential applications of a successful face recognition system. The environment surrounding a face recognition application can cover a wide spectrum – from a well controlled environment to an uncontrolled one. In a controlled environment, frontal and profile photographs of human faces are taken, complete with a uniform background and identical poses among the participants. [16] These face images are commonly called mug shots. Each mug shot can be manually or automatically cropped to extract a normalized subpart called a canonical face image, as shown in Fig. In a canonical face image, the size and position of the face are normalized approximately to the predefined values and the

background region is minimized. Face recognition techniques for canonical images have been successfully developed by many face recognition systems.



Figure 3: A few examples of canonical frontal face images.

General face recognition, a task which is done by humans in daily activities, comes from a virtually uncontrolled environment. Systems to automatically recognize faces from uncontrolled environment must first detect faces in sensed images. A scene may or may not contain a set of faces; if it does, their locations and sizes in the image must be estimated before recognition can take place by a system that can recognize only canonical faces. A face detection task is to report the location, and typically also the size, of all the faces from a given image. Figure 3. gives an example of an image which contains a number of faces. From figure 3, we can see that recognition of human faces from an uncontrolled environment is a very complex problem, more than one face may appear in an image; lighting condition may vary tremendously; facial expressions also vary from time to time; faces may appear at different scales, positions and orientations; facial hair, make-up and turbans all obscure facial features which may be useful in localizing and recognizing faces; and a face can be partially occluded. [5], [23], [39] Further, depending on the application, handling facial features over time (e.g., aging) may also be required. Given a face image to be recognized, the number of individuals to be matched against is an important issue. [11] This brings up the notion of face recognition versus verification: given a face image, a recognition system must provide the correct label (e.g., name label) associated with that face from all the individuals in its database. A face verification system just decides if an input face image is associated with a given face image. Since face recognition in a general setting is very difficult, an application system typically restricts one of many aspects, including the environment in which the recognition system will take place (fixed location, fixed lighting, uniform background, single face, etc.), the allowable face change (neutral expression, negligible aging, etc.), the number of individuals to

be matched against, and the viewing condition (front view, no occlusion, etc.).



Figure 4: An image that contains a number of faces.

The task of face detection is to determine the position and size (height and width) of a frame in which a face is canonical. Such a frame for a particular face is marked in the image.[15]

III. FACE DETECTION

Face Detection is a part of a wide area of pattern Detection technology. Detection and especially face Detection covers a range of activities from many walks of life. Face Detection is something that humans are particularly good at and science and technology have brought many similar tasks to us. Face Detection in general and the Detection of moving people in natural scenes in particular, require a set of visual tasks to be performed robustly. That process includes mainly three-task acquisition, normalisation and Detection. By the term acquisition we mean the detection and tracking of face-like image patches in a dynamic scene. Normalisation is the segmentation, alignment and normalisation of the face images[3], and finally Detection that is the representation and modelling of face images as identities, and the association of novel face images with known models.

IV. Principal Component Analysis

On the field of face Detection most of the common methods employ Principal Component Analysis. Principal Component Analysis is based on the Karhunen-Loeve (K-L), or Hostelling Transform, which is the optimal linear method for[9] reducing redundancy, in the least mean squared

reconstruction error sense. 1. PCA became popular for face Detection with the success of eigenfaces.

The idea of principal component analysis is based on the identification of linear transformation of the co-ordinates of a system. "The three axes of the new co-ordinate system coincide with the directions of the three largest spreads of the point distributions."

In the new co-ordinate system that we have now the data is uncorrected with the data we had in the first co-ordinate system. [2]

For face Detection, given dataset of N training images, we create N d -dimensional vectors, where each pixel is a unique dimension. The principal components of this set of vectors is computed in order to obtain a $d \times m$ projection matrix, W . Approximates the original image where μ is the mean, of the χ_i and the reconstruction is perfect when $m = d$.

For the comparison we are going to use two different PCA algorithms. The first algorithm[11] is computing and storing the weight of vectors for each person's image in the training set, so the actual training data is not necessary. In the second algorithm each weight of each image is stored individually, is a memory-based algorithm. For that we need more storing space but the performance is better.

In order to implement the Principal component analysis in MATLAB we simply have to use the command `prepca`. The syntax of the command is

```
ptrans,transMat = prepca(P,min_frac)
```

`Prepca` pre-processes the network input training set by applying a principal component analysis. This analysis transforms the input data so that the elements of the input vector set will be uncorrected. In addition, the size of the input vectors may be reduced by retaining[10] only those components, which contribute more than a specified fraction (`min_frac`) of the total variation in the data set.

`Prepca` takes these inputs the matrix of centred input (column) vectors, the minimum fraction variance component to keep and as result returns the transformed data set and the transformation matrix.

a) Algorithm

Principal component analysis uses singular value decomposition to compute the principal components. A matrix whose rows consist of the eigenvectors of the input covariance matrix

multiplies the input vectors. This produces transformed input vectors whose components are uncorrected and ordered according to the magnitude of their variance.

Those components, which contribute only a small amount to the total variance in the data set, are eliminated. It is assumed that the input data set has already been normalised so that it has a zero mean.

In our test we are going to use two different “versions” of PCA. In the first one the centroid of the weight vectors for each person’s images in the training set is computed and stored. On the other hand in PCA-2 a memory based variant of PCA, each of the weight vectors in individually computed and stored.

Eigenfaces

Human face Detection is a very difficult and practical problem in the field of pattern Detection. On the foundation of the analysis of the present methods on human face Detection, [12] a new technique of image feature extraction is presented. And combined with the artificial neural network, a new method on human face Detection is brought up. By extraction the sample pattern's algebraic feature, the human face image's eigenvalues, the neural network classifier is trained for Detection. The Kohonen network we adopted can adaptively modify its bottom up weights in the course of learning. Experimental results show that this method not only utilises the feature aspect of eigenvalues but also has the learning ability of neural network. It has better discriminate ability compared with the nearest classifier. The method this paper focused on has wide application area. The adaptive neural network classifier can be used in other tasks of pattern Detection.

In order to calculate the eigenfaces and eigenvalues in MATLAB we have to use the command eig. The syntax of the command is

$d = \text{eig}(A)$

$V, D = \text{eig}(A)$

$V, D = \text{eig}(A, 'nobalance')$

$d = \text{eig}(A, B)$

$V, D = \text{eig}(A, B)$

$d = \text{eig}(A)$ returns a vector of the eigenvalues of matrix A. $V, D = \text{eig}(A)$ produces matrices of

eigenvalues (D) and eigenvectors (V) of [13] matrix A, so that $A*V = V*D$. Matrix D is the canonical form of A, a diagonal matrix with A's eigenvalues on the main diagonal. Matrix V is the modal matrix, its columns are the eigenvectors of A. The eigenvectors are scaled so that the norm of each is 1.0. Then we use $W, D = \text{eig}(A')$; $W = W'$ in order to compute the left eigenvectors, which satisfy $W*A = D*W$.

$V, D = \text{eig}(A, 'nobalance')$ finds eigenvalues and eigenvectors without a preliminary balancing step. Ordinarily, balancing improves the conditioning of the input matrix, enabling more accurate computation of the eigenvectors and eigenvalues. However, if a matrix contains small elements that are really due to round-off error, balancing may scale them up to make them as significant as the other elements of the original matrix, leading to incorrect eigenvectors. We can use the no balance option in this event.

$d = \text{eig}(A, B)$ returns a vector containing the generalised eigenvalues, if A and B are square matrices. $V, D = \text{eig}(A, B)$ produces a diagonal matrix D of generalised eigenvalues and a full matrix V whose columns are the corresponding eigenvectors so that $A*V = B*V*D$. The eigenvectors are scaled so that the norm of each is 1.0.

Euclidean distance

One of the ideas on which face Detection is based is the distance measures, between to points. The problem of finding the distance between two or more point of a set is defined as the Euclidean distance. The Euclidean distance is usually referred to the closest distance between two or more points.

IV. IMPLEMENTATION

The first component of our system is a filter that receives as input a 20x20 pixel region of the image, and generates an output ranging from 1 to -1, signifying the presence or absence of a face, respectively. To detect faces anywhere in the input, the filter is applied at every location in the image. To detect faces larger than the window size, the input image is repeatedly reduced in size (by subsampling), and the filter is applied at each size. This filter must have some invariance to position and scale. The amount of invariance determines the number of scales and positions at which it must be applied. For the work presented here, we apply the filter at every pixel position in the image, and scale

the image down by a factor of 1.2 for each step in the pyramid. The filtering algorithm is shown in . First, a preprocessing step, adapted from , is applied to a window of the image. The window is then passed through a neural network, which decides whether the window contains a face. The preprocessing first attempts to equalize the intensity values in across the window. We fit a function which varies linearly across the window to the intensity values in an oval region inside the window. Pixels outside the oval may represent the background, so those intensity values are ignored in computing the lighting variation across the face. The linear function will approximate the overall brightness of each part of the window, and can be subtracted from the window to compensate for a variety of lighting conditions. Then histogram equalization is performed, which non-linearly maps the intensity values to expand the range of intensities in the window. The histogram is computed for pixels inside an oval region in the window. This compensates for differences in camera input gains, as well as improving contrast in some cases. For the experiments which are described later, we use networks with two and three sets of these hidden units. Similar input connection patterns are commonly used in speech and character recognition tasks .The network has a single, real-valued output, which indicates whether or not the window contains a face. The network has some invariance to position and scale, which results in multiple boxes around some faces. To train the [14]neural network used in stage one to serve as an accurate filter, a large number of face and nonface images are needed. Nearly 1050 face examples were gathered fromface databases at CMU, Harvard2, and from the World Wide Web. The images contained faces of various sizes, orientations, positions, and intensities. The eyes, tip of nose, and corners and center of the mouth of each face were labelled manually. These points were used to normalize each face to the same scale, orientation, and position, as follows:

Table 1: Methodology

- a.) Use LDA and Fishers Face Algorithm.
- b.) Take Training data base.
- c.) Take Test image.
- d.) Implementation of the PCA and LDA.
- e.) Checking the test image on training data.
- f.) Compilation and Performance graph generation on the ease of steps b, c, d, and e.

Now the algorithm for the proposed technique is as follows:

```
Step1. Align a set of face images say T
Step 2. Create training database (ORL Face
database) of M rows and N columns of each image.
P=M x N
Step3. Reshapes: 2D images into 1D column
vectors.
Step 4. Create database

W=26           % number of folders in database
for i=1: w      %for each unit of database

    if DB=1 Then % where DB is the database means
database exists
    DB= 1: i
    Find Components
    Ti is mapped onto a (P-C) mapping
    if  $D_{min} == 0$  then %where  $D_{min}$  is the minimum
value of the %mean distance between test image
and trained image
    Proceed
    Else
    Goto step 4 again;
    Endif
    End For
Step 5. Calculating Discriminant for Fisher Linear
(P-C)(C-1)
    for DB=1: w
    Projected Images Fisher
    for 1: (C-1)*P
    %Training images from 1 to w
    End for
    End for
Show the Matched Output with Success rate
```

V. RESULTS

The database of images is having the images of 10 different peoples and we are performing our test on 3 of them. The following results were found.

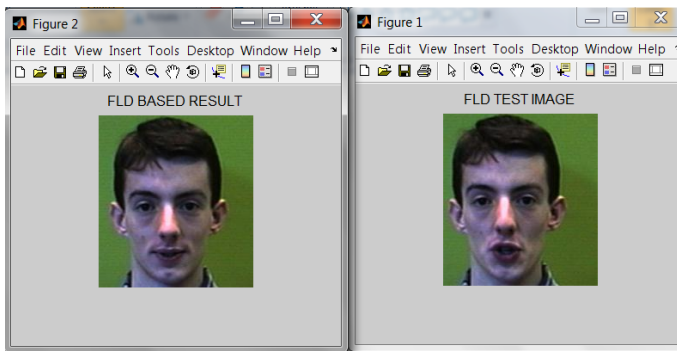


Figure 6: Test image for FLD testing (image 1/10).

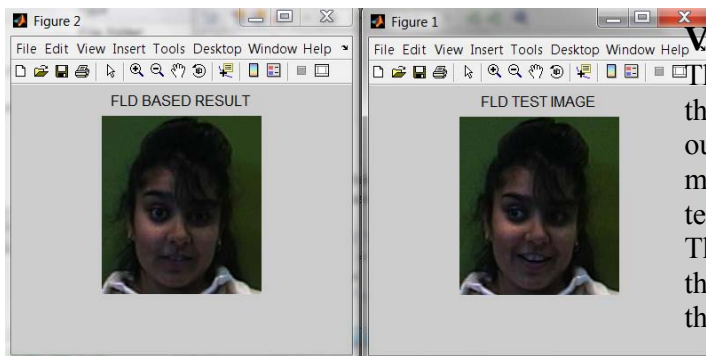


Figure 7: Test image for FLD testing (image 2/10).



Figure 8: Test image for FLD testing (image 3/10).

Approach	No. of correct outputs out of 100	Accuracy Rate (%)
PCA	90	90
Proposed PCA along with linear distance finding method	99	99

The ORL Database of Facial Images [19] is used for performing the experiments. The database consists of 400 facial images of 40 individuals with 10 images of each. For performing the experiments we have taken 100 images of 10 individuals with 10 images of each. The training set consists of 50 images from these with 5 images of each individual.

The experiment is performed first by recognizing images of each individual using PCA and then PCA with linear distance finding algorithm. Then, the accuracy rate for both the approaches is calculated, by finding out, how many results are found correct.

VI. Conclusion

The propose work shows the robust performance for the give test images the achieved output is 99% in our case. The system performance may vary machine to machine. In our system, we perform the test on i3 machine with 4GB Ram in less than 5 sec. The speed performance and accuracy outperforms the available methods till date. Our system is better than the all available methods of face recognition.

VII. REFERENCES

1. "Face recognition using pca, lda and ica approaches on colored images:.", Önsen Toygar Adnan, Acan, Journal Of Electrical & Electronics Engineering Year 2003 Volume 3 pp(735-743).
2. "A modified Hausdorff distance for object matching:." M.P. Dubuisson and A.K. Jain, In ICPR94, pages A:566–568, Jerusalem, Israel, 1994.
3. "The extended M2VTS database:." K. Messer, J. Matas, J. Kittler, J. Luetttin, and G. Maitre, In Second International Conference on Audio and Video-based Biometric Person Authentication, pages 72–77, March 1999.
4. "Hierarchical discriminant analysis for image retrieval:." D. Swets and J. Weng, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 21, issue 5 pp 386–401, 1999.
5. "Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition

- Algorithms.” Wendy S. Yambor, July 2000 (Technical Report CS-00-103, Computer Science).
6. “Face Recognition using Principle Component Analysis.” Kyungnam Kim Department of Computer Science University of Maryland, College Park MD 20742, USA 2003.
 7. “Face Recognition Using Eigenfaces.” M.A. Turk and A.P. Pentland, IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586-591, 1991.
 8. “Principal Component Neural Networks: Theory and Applications.” K. I. Diamantaras and S. Y. Kung, John Wiley & Sons, Inc., 1996.
 9. “View-Based and Modular Eigenspaces for Face Recognition.” Alex Pentland, Baback Moghaddam, and Thad Starner, IEEE Conf. on Computer Vision and Pattern Recognition, MIT Media Laboratory Tech. Report No. 245 1994
 10. “Mechanisms of human facial recognition.” R. Baron, International Journal of Man-Machine Studies, pp. 137-178, 1981.
 11. “Familiarity and recognition of faces in old age.” J. C. Bartlett and A. Fulton, Memory and Cognition, Vol. 19, No. 3, pp. 229-238, 1991.
 12. “Eigenfaces vs fisherfaces: Recognition using class specific linear projection.” P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711-720, 1997.
 13. “Survey: Face recognition systems.” C. Bunney, Biometric Technology Today, Vol. 5, No. 4, pp. 8-12, 1997.
 14. “Individual face classification by computer vision.” R. A. Campbell, S. Cannon, G. Jones, and N. Morgan, In Proceedings Conference Modeling Simulation Microcomputer, pp. 62-63, 1987.
 15. “A case study: Face recognition.” S. Carey, In Explorations in the Biological Language. Bradford Books, New York, 1987.
 16. “The development of face recognition-a maturational component?” S. Carey, R. Diamond, and B. Woods, Developmental Psychology, No. 16, pp. 257-269, 1980.
 17. “Human and machine recognition of faces: A survey.” R. Chellappa, C. L. Wilson, and S. Sirohey, Proceedings of the IEEE, Vol. 83, No. , pp. 705-740, 1995.
 18. “Introduction to aspects of face processing: Ten questions in need of answers.” H. D. Ellis, In H. Ellis, M. Jeeves, F. Newcombe, and A. Young, editors, Aspects of Face Processing, pp. 3-13. Nijhoff, 1996.
 19. “Priming effects in children's face recognition.” H. D. Ellis, D. M. Ellis, and J. A. Hosie, British Journal of Psychology, Vol. 84, No. 1, pp. 101-110, 1993.
 20. “Rethinking Innateness: A connectionist perspective on development.” J. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett, MIT Press, Cambridge, MA, 1997.
 21. “Discriminant analysis for recognition of human face images.” K. Etemad and R. Chellappa, In Proceedings International Conference Acoustics, Speech, Signal Processing, pp. 2148-2151, Atlanta, Georgia, 1994.
 22. FG1, Proceedings of International Workshop on Automatic Face- and Gesture-Recognition. Multimedia lab. Department of Computer Science, University of Zurich, Zurich, Switzerland, 1995.
 23. FG2, Proc. 2nd International Conference on Automatic Face and Gesture Recognition. IEEE Computer Society Press, Los Alamitos, CA, 1996.
 24. FG3, Proceedings 3rd International Conference on Automatic Face and Gesture Recognition. IEEE Computer Society Press, Los Alamitos, CA, 1998.
 25. “Facial feature variation: Anthropometric data II.” A. G. Goldstein, Bulletin of the Psychonomic Society, Vol. 13, pp. 191-193, 1979.
 26. “Race-related variation of facial features: Anthropometric data I.” A. G. Goldstein, bulletin of the Psychonomic Society, Vol. 13, pp. 187-190, 1979.
 27. “Biology and cognitive development: The case of face recognition.” P. Green, Animal Behaviour, Vol. 43, No. 3, pp. 526-527, 1992.
 28. “The human face.” D. C. Hay and A. W. Young, In H. D. Ellis, editor, Normality and Pathology in Cognitive Function, pp. 173-202. Academic Press, London, 1982.

29. "Computer Recognition of Human Faces:" T. Kanade, Birkhauser, Basel and Stuttgart, 1977.
30. "A basic study on human face recognition:" Y. Kaya and K. Kobayashi, In Wantanabe S., editor, *Frontiers of Pattern Recognition*, pp. 265-289. Academic Press, New York, 1972.
31. "Application of the karhunen-love procedure for the characterization of human faces:" M. Kirby and L. Sirovich, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 103-108.
32. "Self-Organization and Associative Memory:" T. Kohonen, Springer-Verlag, Berlin, 1988.
33. "Face recognition: A convolutionalnet work approach:" S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, *IEEE Trans. Neural Networks*, Vol. 8, No. 1, pp. 98-113, 1997.
34. "Caricature and face recognition:" R. Mauro and M. Kubovy, *Memory and Cognition*, Vol. 20, No. 4, pp. 433-441, 1992.
35. "Recognizing and naming faces: aging, memory retrieval, and the tip of the tongue state:" E. A. Maylor, *Journal of Gerontology*, Vol. 45, No. 6, pp. 215-226, 1990.
36. "Conspec and conlern: a two-process theory of infant face recognition:" J. Morton and M. H. Johnson, *Psychological Review*, Vol. 98, No. 2, pp. 164-181, 1991.
37. "Visual learning and recognition of 3-D objects from appearance:" H. Murase and S. K. Nayar, *Internationa Journal of Computer Vision*, Vol. 14, No. 1, pp. 5-24, 1995.
38. "The FERET evaluation methodology or face-recognition algorithms:" P. J. Phillips, H. Moon, P. Rauss, and S. A. Rizvi, In *Proceedings IEEE Conf. Computer Vision and Pattern Recognition*, pp. 137-143, Puerto Rico, 1997.
39. "To catch a thief with a recognition test: the model and some empirical results:" S. S. Rakover and B. Cahlon, *Cognitive Psychology*, Vol. 21, No. 4, 423-468, 1989.
40. "The simon then garfunkel effect: semantic priming, sensitivity, and the modularity of face recognition:" G. Rhodes and T. Tremewan, *Cognitive Psychology*, Vol. 25, No. 2, pp. 147-187, 1993.
41. "Neural network-based face detection:" H. A. Rowley, S. Baluja, and T. Kanade, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, 23-38, 1998.
42. "Face recognition/detection by probabilistic decision-based neural network:" S. K. S. Lin and L. Lin, *IEEE Trans. Neural Networks*, Vol. 8, No. 1, pp. 114-132, 1997.
43. "Automatic recognition and analysis of human faces and facial expressions: A survey:" Samal and P. Iyengar, *Pattern Recognition*, Vol. 25, pp. 65-77, 1992.
44. "Microgenesis of face perception:" J. Sergeant, In *Aspects of Face Processing*. Nijhoff, Dordrecht, 1986.
45. "Visual-field asymmetry in face recognition as a function of face discriminability and interstimulus interval:" W. Sjoberg, B. Gruber, and C. Swatloski, *Perceptual and Motor Skills*, Vol. 72, No. 3, pp. 1267-1271, 1991.
46. "Example-based learning for view-based human face detection:" K. K. Sung and T. Poggio, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, pp. 39-51, 1998.
47. "PCA versus LDA:" A. M. Martinez and A. C. Kak, *IEEE Trans. On pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pp. 228-233, 2001.
48. "Automatic Face recognition using neural network- PCA:" Boualleg, A.H., Bencheriet, Ch., Tebbikh, *Information and Communication Technologies*, 2006. ICTTA '06. 2nd Volume 1, 24-28 April 2006
49. "Face recognition by using neural network classifiers based on PCA and LDA:" Byung-Joo Oh, *Systems, man & Cybernetics*, 2005 IEEE international conference.
50. "Personal identification and description:" Francis Galton, In *Nature*, pp. 173-177, June 21, 1888.
51. "Robust image based 3D face recognition:" W. Zaho, Ph.D. Thesis, Maryland University, 1999.
52. "Human and machine recognition of faces: A survey:" R. Chellappa, C. L. Wilson, and S. Sirohey, *Proc. IEEE*, vol. 83, pp. 705-741, May 1995.
53. "Discriminant analysis and eigenspace partition tree for face and object recognition from views:" D. Swets and J. Weng, In *Proc. Int'l Conference on Automatic Face-*

- and *Gesture-Recognition*, pages 192-197, Killington, Vermont, 1996.
54. "Using discriminant eigenfeatures for image retrieval." D. L. Swets and J. Weng, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, pp. 831-836, 1996.
 55. "Accustomed to your face:" M. Szpir, *American Scientist*, Vol. 80, No. 6, 537-540, 1992.
 56. "Margaret Thatcher: a new illusion:" P. Thompson, *Perception*, Vol. 9, pp. 483-484, 1980.
 57. "Eigenfaces for recognition:" M. Turk and A. Pentland, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
 58. "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis : " Hyvärinen, A., *IEEE Transactions on Neural Networks*, Vol. 10, No. 3, pp. 626-634, 1999.
 59. "Description of Libor Spacek's Collection of Facial Images:" Spacek L., 1996, online: <http://cswww.essex.ac.uk/allfaces/index.htm> l.
 60. "An Efficient LDA Algorithm for Face Recognition:" Yang J., Yu Y., Kunz W., The Sixth International Conference on Control, Automation, Robotics and Vision (ICARCV2000), 2000.
 61. "PCA versus LDA:" Martinez A.M. and Kak A.C., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No.2, pp. 228-233, 2001.
 62. "Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms:" Yambor W.S., Technical Report CS-00-103, Computer Science Department, Colorado State University, July 2000.
 63. "A Survey of Face Recognition:" Fromherz T., Stucki P., Bichsel M., MML Technical Report, No 97.01, Dept. of Computer Science, University of Zurich, 1997.
 64. "Typicality in categorization, recognition and identification: evidence from face recognition:" T. Valentine and A. Ferrara, *British Journal of Psychology*, Vol. 82, No. 1), pp. 87-102, 1991.
 65. "Prior knowledge and face recognition in a community-based sample of healthy, very old adults:" A. Wahlin, L. Backman, T. Mantyla, A. Herlitz, M. Viitanen, and B. Winbald, *Journals of Gerontology*, Vol. 48, No. 2, pp. 54, 1993.
 66. "On comprehensive visual learning:" J. Weng, *In Proceedings NSF/ARPA Workshop on Performance vs. Methodology in Computer Vision*, pp. 152-166, Seattle, WA, 1994.
 67. "Learning recognition and segmentation using the Cresceptron:" J. Weng, N. Ahuja, and T. S. Huang, *In Proceedings International Conference on Computer Vision*, pp. 121-128. Berlin, Germany, 1993.
 68. "Pca based face recognition and testing criteria:" Bruce Poon, M. Ashraful Amin, Hong Yan, *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics*, Baoding, 12-15 July 2009.
 69. "Research on Face Recognition Based on PCA:" Hong Duan1, Ruohe Yan1, Kunhui Lin, 2008 International Seminar on Future Information Technology and Management Engineering.
 70. "New Parallel Models for Face Recognition:" Heng Fui Liao, Kah Phooi Seng, Yee Wan Wong, Li-Minn Ang, 2007 International Conference on Computational Intelligence and Security.
 71. "Median LDA: A Robust Feature Extraction Method for Face Recognition:" Jian Yang, David Zhana and Jing-yu Yang, 2006 IEEE International
 72. "An improved WPCA plus LDA:" Bai Xiaoman, Ruan Qiuqi, ICSP2006 Proceedings.
 73. "Is ICA Significantly Better than PCA for Face Recognition? : " Jian Yang David Zhang, Jing-yu Yang, *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)*.
 74. "On the Euclidean Distance of Images : " Liwei Wang, Yan Zhang, and Jufu Feng, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol. 27, No. 8, August 2005.
 75. "Face Recognition Using Laplacianfaces : " Xiaofei He, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, and Hong-Jiang Zhang, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol. 27, No. 3, March 2005.

76. "A modified pca algorithm for face recognition:" Lin Luo, M.N.S. Swamy, Eugene I. Plotkin, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 12, No. 1, pp. 57-60, 1999.
77. "Face Recognition Using LDA Mixture Model:" Hyun-Chul Kim, Daijin Kim, Sung Yang Bang, IEEE Transactions On Neural Networks, Vol. 8, No. 1, January 2002.
78. "Discriminant Analysis of Principal Components for Face Recognition:" W. Zhao R. Chellappa A. Krishnaswamy, IEEE Transactions On Neural Networks, Vol. 8, No. 1, January 1997.
79. "Hierarchical Discriminant Analysis for Image Retrieval:" Daniel L. Swets, and Juyang Weng, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 21, No. 5, May 1999.
80. "Neural Network-Based Face Detection:" Henry A. Rowley, Shumeet Baluja, and Takeo Kanade, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 20, No. 1, January 1998.
81. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection:" Peter N. Belhumeur, Jo~ao P. Hespanha, and David J. Kriegman, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 19, No. 7, July 1997.
82. "Face Recognition/Detection by Probabilistic Decision-Based Neural Network:" Shang-Hung Lin, Sun-Yuan Kung, and Long-Ji Lin, IEEE Transactions On Neural Networks, Vol. 8, No. 1, January 1997.
83. "Face Recognition: A Convolutional Neural-Network Approach:" Steve Lawrence C. Lee Giles, Ah Chung Tsoi, and Andrew D. Back, IEEE Transactions On Neural Networks, Vol. 8, No. 1, January 1997.
84. "Human and Machine Recognition of Faces: A Survey:" Rama Chellappa, Charles L. Wilson, Proceedings Of The IEEE, Vol 83, No 5, May 1995.
85. "Beyond Eigenfaces: Probabilistic Matching for Face Recognition:" B. Moghaddam, W. Wahid, and A. Pentland, The 3rd International Conference on Automatic Face & Gesture Recognition, Nara, Japan, April 1998.
86. "Probabilistic Visual Learning for Object Representation:" B. Moghaddam, and A. Pentland, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, July 1997.
87. "Principal Manifolds and Probabilistic Subspaces for Visual Recognition:" B. Moghaddam, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 6, June 2002.
88. "A Unified Bayesian Framework for Face Recognition:" C. Liu, and H. Wechsler, pp. 151-155, IEEE, 1998.
89. "Probabilistic Reasoning Models for Face Recognition:" C. Liu, and H. Wechsler, pp. 827-832, IEEE, 1998.
90. "Face Recognition using Principal Component Analysis of Gabor Filter Responses:" K. C. Chung, S. C. Kee, and S. R. Kim, pp. 53- 57, IEEE, 1999.
91. "A Local Face Statistics Recognition Methodology beyond ICA and/or PCA:" A. X. Guan, and H. H. Szu, pp. 1016-1027, IEEE, 1999.
92. "Pattern Classification : " R. O. Duda, P. E. Hart, and D. G. Stork, John Wiley & Sons, 2nd Edition, 2001.

Robust & Accurate Face Recognition using Histograms

Sarbjeet Singh¹, Meenakshi Sharma², Dr. N.Suresh Rao³

Mtech CSE(4th sem)¹,HOD CSE²,HOD MCA³

SSCET Pathankot^{1,2},Jammu University³

sarbaish@gmail.com¹, mss.s.c.e.t@gmail.com²

Abstract : *A large number of face recognition algorithms have been developed from decades. Face recognition systems have been grabbing high attention from commercial market point of view as well as pattern recognition field. It also stands high in researchers community. Face recognition have been fast growing, challenging and interesting area in real-time applications. This face recognition system detects the faces in a picture taken by web-cam or a digital camera, and these face images are then checked with training image dataset based on descriptive features. In this paper , we use a histogram approach for human face detection. Since different faces contains different facial features, having the features which are unique. In this paper the vector machine is used for skin detection and face detection.*

Keywords : Face recognition ,PCA, LDA Histogram.

1.Introduction :

Face recognition is one of the most active and widely used technique[1-2] because of its reliability and accuracy in the process of recognizing and verifying a person's identity. The need is becoming important since people are getting aware of security and privacy. For the Researchers Face Recognition is among the tedious work. It is all because the human face is very robust in nature; in fact, a person's face can change very much during short periods of time (from one day to another) and because of long periods of time (a difference of months or years). One problem of face recognition is the fact that different faces could seem very similar; therefore, a discrimination task is needed. On the other hand,

when we analyze the same face, many characteristics may have changed. These changes might be because of changes in the different parameters. The parameters are: illumination, variability in facial expressions, the presence of accessories (glasses, beards, etc); poses, age, finally background. We can divide face recognition[7-8] techniques into two big groups, the applications that required face identification and the ones that need face verification. The difference is that the first one uses a face to match with other one on a database; on the other hand, the verification technique tries to verify a human face from a given sample of that face.

2. Histogram

Histogram, or Frequency Histogram is a bar graph. The horizontal axis depicts the range and scale of observations involved and vertical axis shows the number of data points in various intervals ie. the frequency of observations in the intervals.

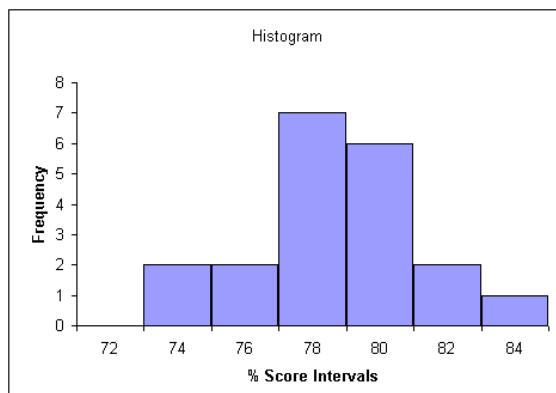
Histograms are popular among statisticians. Though they do not show the exact values of the data points they give a very good idea about the spread of the data and shape.

Let us try drawing a histogram of percentage scores in a test . The scores are as follows :-

82.5, 78.3, 76.2, 81.2, 72.3, 73.2, 76.3, 77.3, 78.2, 78.5, 75.6, 79.2, 78.3, 80.2, 76.4, 77.9, 75.8, 76.5, 77.3, 78.2

When any data is provided to XLMiner[◆], it decides the size and number of intervals amongst which the data should be distributed. It uses "Nicing" to decide the number of intervals. Five to Twenty intervals are fixed on the dataset depending on its range.

Now see the histogram of the same data.



The values on the horizontal axis are the **upper limits** of bins (intervals) of data points, and not the mid-points of the intervals, although they may appear to be so. This is in keeping with the way the Analysis Toolpak of Excel works. As an example, the bar

shown against 78 has a frequency of 7. That means 7 data points lie in the range above 76 and upto (including) 78.

As is evident, the histogram gives a fairly good idea about the shape and spread of data at a glance.

3. Face Recognition

Face recognition is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness. It has the accuracy of a physiological approach without being intrusive. For this reason, since the early 70's, face recognition has drawn the attention of researchers in fields from security, psychology, and image processing, to computer vision. Numerous algorithms have been proposed for face recognition; While network security and access control are its most widely discussed applications, face recognition has also proven useful in other multimedia information processing areas.

Face recognition [5] techniques can be used to browse video database to find out shots of particular people. Also for face images with a compact parameterized facial model for low-bandwidth communication applications such as videophone and teleconferencing. Recently, as the technology has matured, commercial products have appeared on the market. Despite the commercial success of those face recognition products, a few research issues remain to be explored.

3.1 General face recognition system

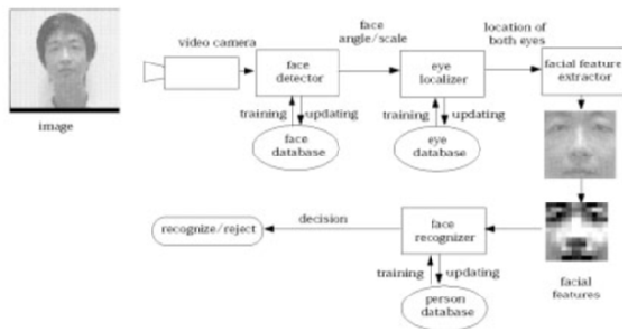


Figure : Block Diagram for Face Recognition System

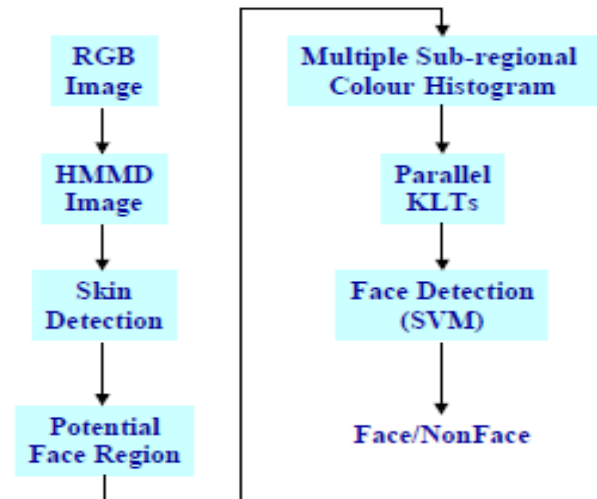


Fig. 2The schematic of the new face recognition/detection method

4. Histogram Method used for Face Detection

As per [9], RGB colour space is commonly used in image processing because of its basic synthesis property and direct application in image display. According to the requirements of different image processing tasks, RGB colour space is often transformed to other colour spaces. From a visual perception's point of view, hue, saturation and value are often employed to manipulate colour, such as desaturation or change of colourfulness. When the colour is quantized to a limit number of representative colours, one will have to deal with two problems. The first is how to best match the distance [3-4] of data representation to human perception. It is desirable that numerical colour distance is proportional to perceptual difference. The second problem is how to best quantize the colours such that the reproductions from these quantized colours is the most faithful to the original. In this work, we adopt a perceptually meaningful colour space, the HMMD colour space, and used a carefully worked out quantization scheme of the MPEG-7 standard

5. Proposed work and Algorithm:

Recognizing objects from large image databases, histogram based methods have proved simplicity and usefulness in last decade. Initially, this idea was based on color histograms. This algorithm presents the first part of our proposed technique named as "Histogram processed Face Recognition" as compared to detection use in [9]

Histogram techniques are well designed for face detection [6] as shown above. But in our case we apply histogram calculation for face recognition. The algorithm given below worked for face recognition with success rate of 95%.

For training, grayscale images with 256 gray levels are used. Firstly, frequency of every gray-level is computed and stored in vectors for further processing. Secondly, mean of consecutive nine frequencies from the stored vectors is calculated and are stored in another vectors for later use in testing phase.

This mean vector is used for calculating the absolute differences among the mean of trained images and

the test image. Finally the minimum difference found identifies the matched class with test image. Recognition accuracy is of 95 in our case.

6. Experimental Results

The ORL Database of Facial Images [19] is used for performing the experiments. The database consists of 400 facial images of 40 individuals with 10 images of each. For performing the experiments we have taken 100 images of 10 individuals with 10 images of each. The training set consists of 50 images from these with 5 images of each individual.

The experiment is performed first by recognizing images of each individual using HISTOGRAM approach. Then, the accuracy rate for both the approaches is calculated, by finding out, how many results are found correct. Table 1.

Approach	No. of correct outputs out of 100	Accuracy Rate (%)
HISTOGRAM	93	93
HISTOGRAM AND INTENSITY VALUE	98	98

7. Algorithm Steps:

Step 1: Take input image I

Step 2: Test the gray level

For I1=1: N %where N is number of Images

Step3: Compute frequency

For I2=1: N

Step 4: Make frequency vector

For I3=1:M %where M is the dimension of frequency vector and taken as M=9

Step5: Calculate mean or mean difference M_d

M_d =Trained
image –Test image

If $M_d = 0$ then
Matched
Got to
Step 7
Else
%Again
Go to
Endif

check for the next image

step 4

Endfor&Goto step 3

Endfor&goto step 2

Endfor& got to step 6

Step 6: Print Not Matched & Stop

Step 7: Show the Mapped Output in GUI & Stop

8. Conclusion :

In this paper, we investigated the use of the Histogram approach and the Histogram approach using intensity values for recognizing images. We compared both the approaches and from the outputs, it was found that for about 50% of individuals, the output image from both the approaches were different, which clearly shows the variation between the two approaches..

Also, it was found from the accuracy rate that the Histogram with pixel intensity value is more accurate as compared to the Histogram only. Hence, Histogram with pixel intensity value approach is recommended for better results in Face Recognition as compared to alone Histogram .

9. Results. Here we are showing outputs for each individual one by one from both the approaches by taking one image for each individual.

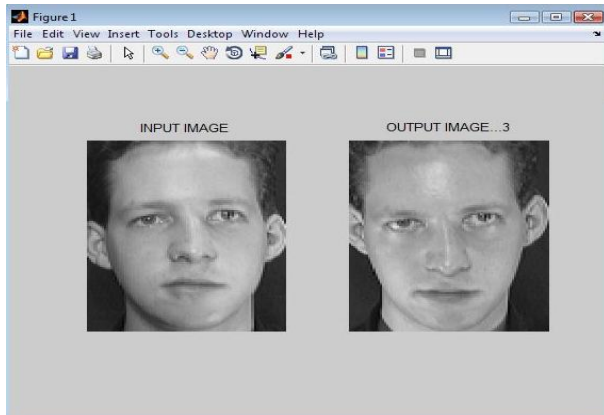


Figure. 9.2. HISTOGRAM Output for First Individual



Figure. 9.4. HISTOGRAM and PIXEL INTENSITY Output for Second Individual



Figure. 9.3. HISTOGRAM Output for Second Individual

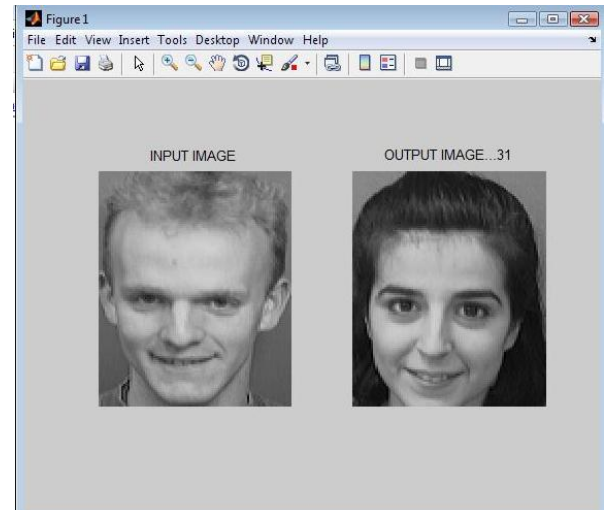


Figure. 9.5. HISTOGRAM Output for Third Individual

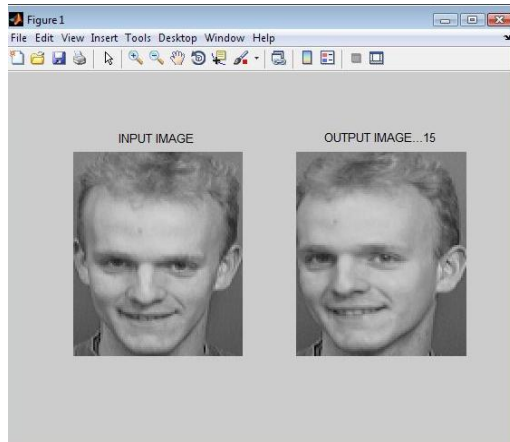


Figure. 9.6. HISTOGRAM and PIXEL INTENSITY Output for Third Individual



Figure. 9.8. HISTOGRAM and PIXEL INTENSITY Output for Fourth Individual



Figure. 9.7. HISTOGRAM Output for Fourth Individual

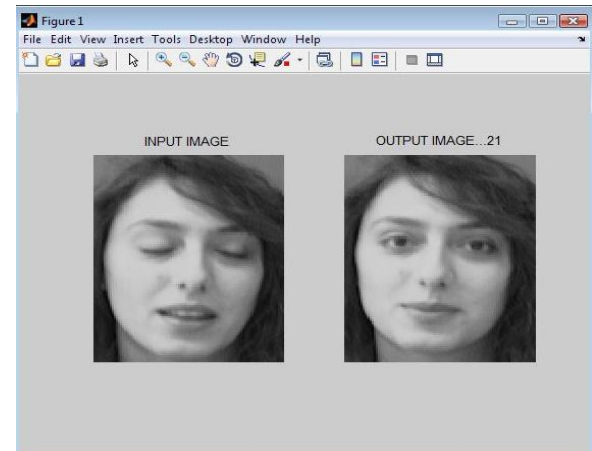
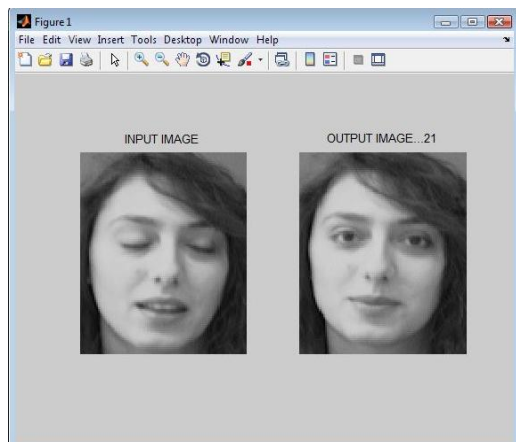


Figure. 9.9. HISTOGRAM Output for Fifth Individual



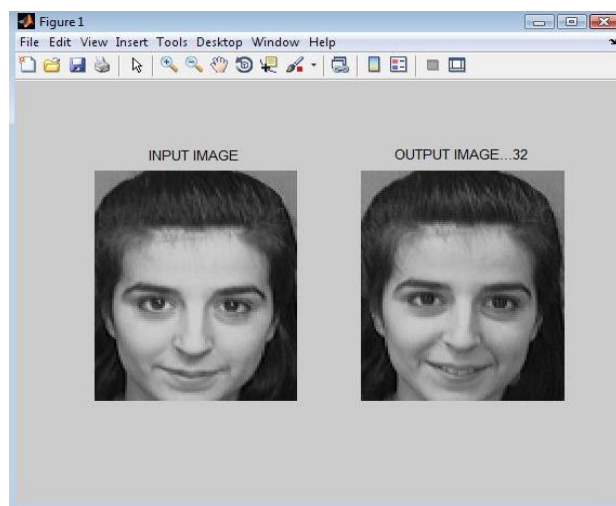
**Figure. 9.10. . HISTOGRAM and PIXEL
INTENSITY Output for Fifth Individual**



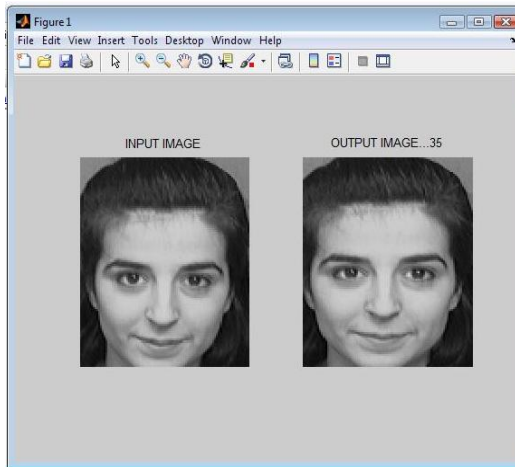
**Figure. 9.12. . HISTOGRAM and PIXEL
INTENSITY Output for Sixth Individual**



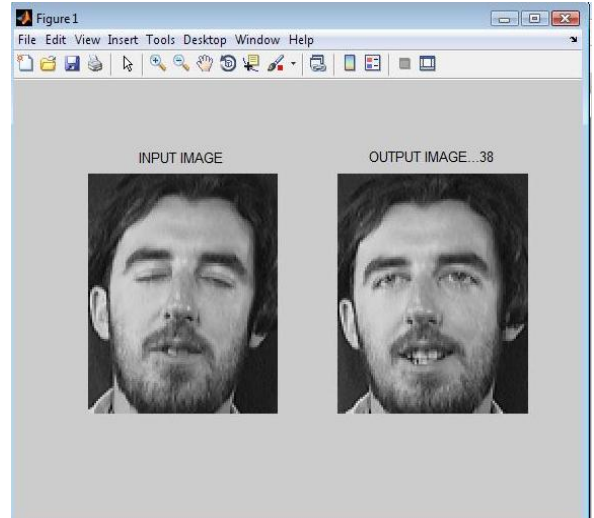
**Figure. 9.11. HISTOGRAM Output for Sixth
Individual**



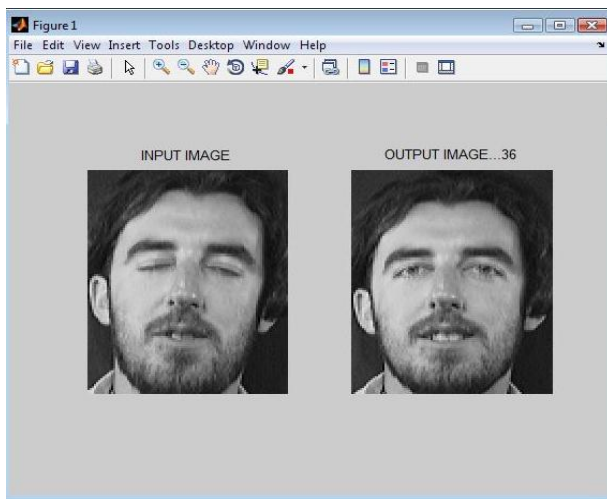
**Figure. 9.13. HISTOGRAM Output for
Seventh Individual**



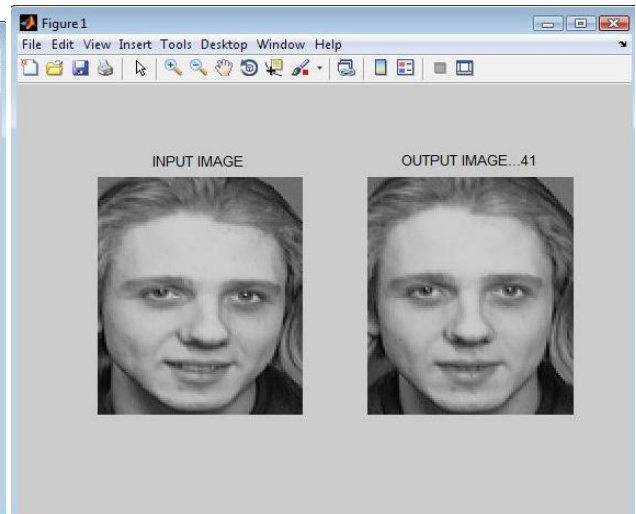
**Figure. 9.14. . HISTOGRAM and PIXEL
INTENSITY Output for Seventh Individual**



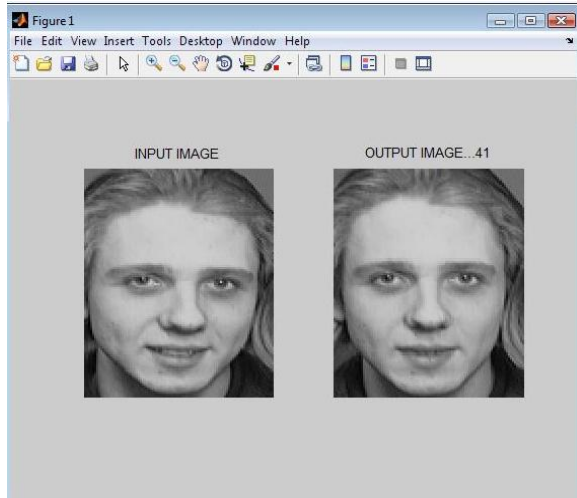
**Figure. 9.16. . HISTOGRAM and PIXEL
INTENSITY Output for Eighth Individual**



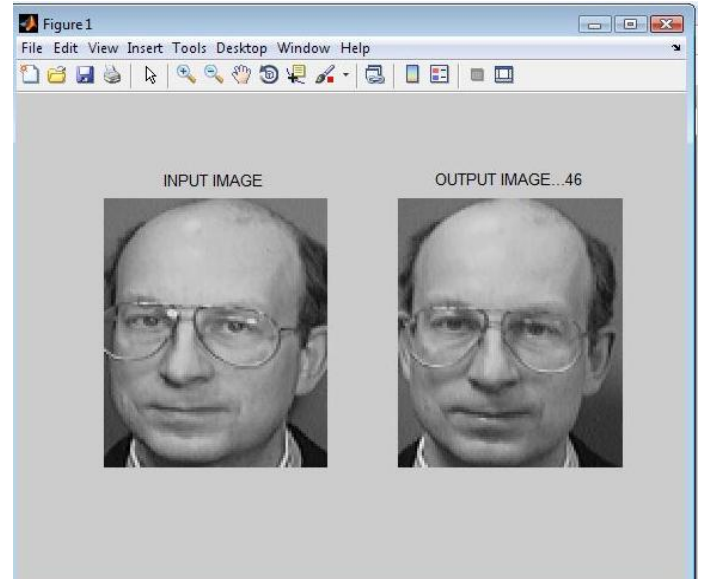
**Figure. 9.15. HISTOGRAM Output for Eighth
Individual**



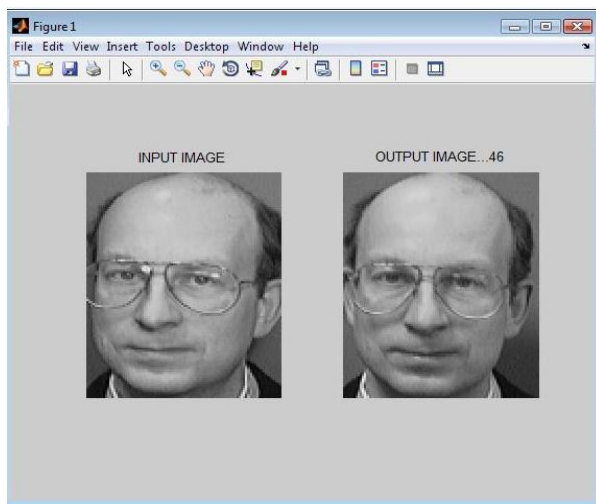
**Figure. 9.17. HISTOGRAM Output for Ninth
Individual**



**Figure. 9.18. HISTOGRAM and PIXEL INTENSITY
Output for Ninth Individual**



**Figure.9.20. . HISTOGRAM and PIXEL
INTENSITY Output for Tenth Individual**



**Figure. 9.19. HISTOGRAM Output for Tenth
Individual**

10. References.

- [1] A. M. Martinez and A. C. Kak, "PCA versus LDA," IEEE Trans. On pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp. 228-233, 2001.
- [2] Boualleg, A.H.; Bencheriet, Ch.; Tebbikh, H "Automatic Face recognition using neural network-PCA" Information and Communication Technologies, 2006. ICTTA '06. 2nd Volume 1, 24-28 April 2006
- [3] Byung-Joo Oh "Face recognition by using neural network classifiers based on PCA and LDA" Systems, man & Cybernetics, 2005 IEEE international conference.
- [4] Francis Galton, "Personal identification and description," In Nature, pp. 173-177, June 21, 1888.
- [5] W. Zaho, "Robust image based 3D face recognition," Ph.D. Thesis, Maryland University, 1999.
- [6] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," Proc. IEEE, vol. 83, pp. 705-741, May 1995.
- [7] T. Riklin-Raviv and A. Shashua, "The Quotient image: Class based recognition and synthesis under varying illumination conditions," In CVPR, P. II: pp. 566-571, 1999.
- [8] G.j. Edwards, T.f. Cootes and C.J. Taylor, "Face recognition using active appearance models," In ECCV, 1998.
- [9] A COLOUR HISTOGRAM BASED APPROACH TO HUMAN FACE DETECTION Jianzhong Fang and Guoping Qiu School of Computer Science, The University of Nottingham

X.509 Authentication Services to Enhance the Data Security in Cloud Computing

Surbhi Chauhan

Department of CSE

Amity University

Noida, INDIA

Surbhichauhan2009@gmail.com

Kamal Kant

Department of CSE

Amity University

Noida, India

kamalkant25@gmail.com

Arjun Singh

Department of CSE

Sir Padampat Singhania University

Udaipur, India

arjun.singh@spsu.ac.in

Abstract— This paper represents a method to build a Cloud Security by giving concept of X.509 authentication services. We are discussing theory of cloud computing, feature of cloud computing and cloud security. We proposed a X.509 format to enhance data security in cloud (Public). Cloud computing is a new computational paradigm that offers an innovative business model for organization.

I. INTRODUCTION

Cloud computing is relay on internet, which have hardware and software base with provision of computing infrastructure. Clouds concept based on existing technologies such as virtualization, utility computing or distributed computing. Cloud computing provides effective IT service delivery and management with efficient lower cost.

A. Service Layers in Cloud Computing

- 1) Software as Service (SaaS): SaaS is at the highest layer and offer application such as service on demand via multitenancy i.e. means a single instance of software serves multiple clients in organization. The example of SaaS is salesforce.com
- 2) Infrastructure as a Service (IaaS)- Cloud outsources the provision of the computing infrastructure which is required to host service. This infrastructure is provided as a service storage and computing resources such as networking, operating system, Load balancers as a cloud service. The high Profile IaaS operation is Amazon's Elastic Compute Cloud (Amazon EC2).
- 3) Platform as a Service (PaaS)- Cloud computing can provide software platform where systems run on execution of services is made in a transparent manner. Clouds systems provide additional abstraction level instead of supplying a virtualized Infrastructure. A well Known example is the Google Apps Engine.

B. Forms of Cloud

Cloud computing can be categories in three types:

1. Private Cloud: Private clouds are on demand infrastructure. It is owned by single customer who controls the application run, and where they have their own servers, networks. Hence the security risk is reduced in Private cloud. Cloud remains behind the firewall to virtualizing the servers.
2. Public Cloud: Public cloud does not depend on any organization; the services provided in Public clouds can be accessed by any organization. Chances of security risk are slightly higher in public cloud.
3. Hybrid Cloud: Hybrid cloud computing is a platform which acts as interface between private cloud and public cloud. It depends on the organizations, which do not want to put everything in the external cloud (public cloud) while we are hosting some servers in their own internal cloud infrastructure.

C. Advantages of Cloud Computing

- i. Faster, simpler and cost effective services
- ii. Highly elastic because resources are occupied on the basis of demand
- iii. Optimized utilization of computing resources
- iv. User virtualizes more resource than they have. For example unlimited storage
- v. Energy efficient as less power consume on hardware and software

D. Securites issues in Cloud

Each type of cloud has certain securities issues. Few securities concern is discussed below.

- i. Many organizations share the resources so there is no absolute control on physical security in cloud model.

- ii. Organization or government can violate the law (risk of data seizure by foreign government)
- iii. Storage services provided by one vendor may be incompatible with another vendor's services, if user wants to move from one vendor to another.
- iv. Ensuring and maintaining the integrity of data is a challenge.
- v. In case of Payment Card Industry Data Security Standard (PCIDSS) data logs must be provided to security managers and regulators.
- vi. User must keep up to date with application improvement to ensure they are protected.
- vii. Due to dynamic and fluid nature of virtual machine, it becomes very difficult to maintain the consistency of security and ensure the auditability of-records.

II. X.509 AUTHENTICATION SERVICE TO ENHANCE DATA SECURITY IN CLOUD

Security is always an issue in cloud computing. In this paper we are proposing X.509 authentication service technique to secure the data in public cloud.

In public cloud there is always a high risk for data, system files, and network traffic and host security as they are vulnerable to attack and has lack of strong authentication mechanism. In this paper, we are proposing the concept of X.509 authentication service to ensure the security of data in cloud. X.509 is relay on asymmetric key cryptography and digital signature. Asymmetric key cryptography and digital signature scheme enhance the security of cloud computing. X.509 technique is widely used in S/MIME IPsec, SSL/TTL and SET.

X.509 has three alternatives authentication procedure, one way authentication, two way authentication and three-way authentications. All these procedures relay on asymmetric key cryptography and digital signature. In asymmetric key cryptography it is assumed that two parties (sender and receiver) share their public key. Here we will apply three way authentication techniques due to its extra advantages over two other procedures.

Let's assume there is two enterprise called A and B as shown in figure 1.0. Enterprise A has public cloud and providing Saas, Paas, IaaS services and Database.

A user from enterprise B wants to access the data in secure manner from the public cloud. Three-way authentications involve transfer of information from A to B in X.509, and establish the following:

- a. A User from enterprise B, sends a request to get a secure data from Enterprise A.
- b. Enterprise A, sends a message consist a nonce $r(a)$, identity of B and message signed with A's private key. The nonce value must be unique and it must be completed within expiration time of message. It is used to detect replay attack.
- c. Enterprise B, sends a message, consist of nonce $r(b)$, identity of enterprise A, sign data with authenticity and integrity, and a session key encrypted with A's public key.

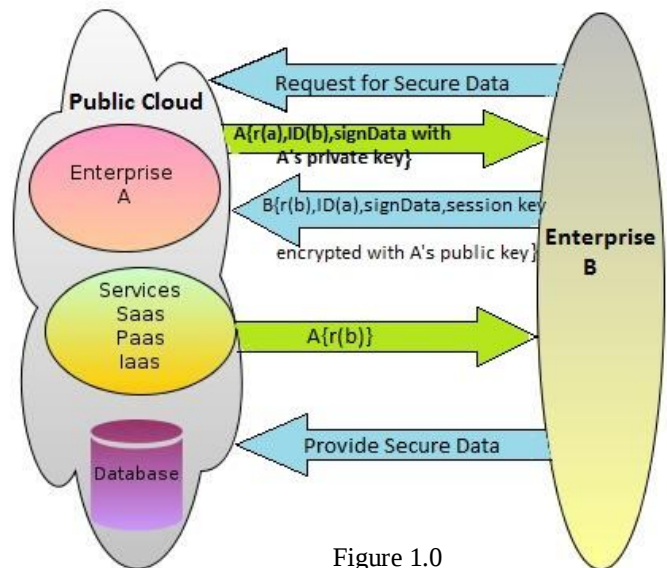


Figure 1.0

- d. A final message from enterprise A to enterprise B sends, which includes a signed copy of the nonce $r(b)$

In three-way authentication, no need to check the timestamp. Each side can check the returned nonce value to detect the replay attacks. On the other hand in two-way authentication, timestamp must be checked.

III. CERTIFICATE

The main part of X.509 is the public key certificate related to each user. These user certificates are created by certification authority (CA). Let's assume, Enterprise A has obtained the certificate from CA, called X1 and enterprise B obtain the certificate called X2. If enterprise A securely knows the public key of X2(Certification Authority), then A can read B's certificate and verifies the signature.

CA signs the certificate (X1) of Enterprise A. User in B must have a copy of the CA's own public Key.

So in cloud computing integrity and authenticity can be enhanced by the X.509 certificate service.

IV. CONCLUSION

As cloud becoming the part of everyone life and mid-size to small-size organizations relaying on cloud, it is essential to secure the data and privacy of transaction done through the cloud computing. In the paper we have discussed a new aspect of implementing existing technology (X.509 service) to enhance the security and integrity of data. X.509 technique also neutralized the replay attack.

IV. REFERENCES

- [1] Amazon Web Services Blog, "Amazon S3, Bigger and Busier Than Ever." [Online]. Available.
<http://aws.typepad.com/aws/2011/01/amazon-s3-bigger-and-busier-than-ever.html>
- [2] Survey by IEEE and Cloud Security Alliance details importance and urgency of Cloud Computing security standards, CSA,
<http://www.cloudsecurityalliance.org/pr20100301c.html>
- [3] Top Threats to Cloud Computing V1.0, CSA, March 2010,
<http://www.cloudsecurityalliance/topthreats.html>
- [4] M. Ouedraogo, H. Mouratidis, D. Khadraoui, and E. Dubois, "An Agent-based System to Support Assurance of Security Requirements," in *Proceedings of the 4th International conference on Secure Software Integration and Reliability Improvement*, 2010.
- [5] S. Ferretti, V. Ghini, F. Panzieri, M. Pellegrini, and E. Turrini, "QoS-aware Cloud," in *Proceedings of the IEEE 3rd International Conference on Cloud Computing*, 2010.
- [6] P. Saripalli and B. Walters, "QUIRC: A Quantitative Impact and Risk Assessment Framework for Cloud Security," in *Proceedings of the IEEE 3rd International Conference of Cloud Computing*, 2010.
- [7] S. A. de Chaves, C. B. Westphall, and F. R. Lamin, "SLA Perspective in Security Management for Cloud Computing" in *Proceedings of the 6th International Conference on Networking and Services*, 2010.

AN INTELLIGENT SPAM-SCAMMER FILTER MECHANISM USING BAYESIAN TECHNIQUES

Olushola D. Adeniji¹, Olubukola Adigun² and Omowumi O. Adeyemo³

^{1,2&3}Department of Computer Science, University of Ibadan, Ibadan, Nigeria

E-Mail: wumiglory@yahoo.com

Abstract

Electronic mail (E-mail) is an electronic message system that transmits messages across computer network. Electronic mail is the easiest and most efficient communication tool for disseminating both wanted and unwanted information. There are many efforts under way to stop the increase of spam that plague almost every user on the internet. Managing and deleting scam or unwanted messages pose negative effects to user's productivity. However the attack of scam on business site also affects the customer. There is an increasing trend of integration of anti-spam techniques into mail transfer agent whereby the mail systems themselves also perform various measures that are generally referred to as filtering, ultimately resulting in spam messages being rejected before delivery (or blocked). This paper presents an E-mail intelligent system using Bayesian algorithm to reduce overload on mail traffic, shutdown of mailbox and waste of disk storage on mail server.

Keywords: *E-mail, Pattern detection, Spam, Filtering, Authentication and Reputation*

1. Introduction

Electronic mail (email) is now considered the easiest and most efficient way to communicate. Internet users can simply type a letter and at the click of a button instantaneously communicate with people all over the world. Electronic mail (E-mail) is an essential communication tool that has been greatly abused by spammers to disseminate unwanted information (messages) and spread malicious contents to Internet users. E-mail serves as an archival tool to some people, while many users never discard messages because their information contents might be useful at a later date as a reminder of upcoming events. The volume and capacity of E-mail that we get is constantly growing. Electronic messages posted blindly to thousands of recipients, and represent one of the most serious and urgent information overload problems. **Lazzari et al. (2005)**. An e-

mail message that is unwanted: Basically it is the electronic version of junk mail that is delivered by the postal service. The term spam refers to unsolicited, unwanted, and inappropriate bulk email. Spam is often referred to as Unsolicited Bulk Email (UBE), Excessive Multi-Posting (EMP), Unsolicited Commercial Email (UCE), and Unsolicited Automated Email (UAE), bulk mail or just junk mail. Spammers use many tactics to get email address to send spam. Another tactics is using social engineering such as chain letter or purchase address from another spammer (Ahmad 2007). They also used computer programs called robots or spiders to harvest email address from websites. Through the internet, spammers can get the email from newsgroup posting, webpage or mailing list. E-mail allows users to communicate with each other at a low cost as well as provides an efficient mail delivery system. The main problem with spam is that it makes up 30% to 60% of mail traffic and is on the rise. It can make the mail traffic become slow. When spam received and storage in mailbox, the mailbox can cause the problem like shutdown.

When dealing with scam, ISP must build a sophisticated program into their system. Other problem at ISP site is server strain[2]. When sending and receiving amount of email in short period of time, server may become strain on ISP resources. They have to upgrade their equipment and pay higher bandwidth bill to deal with the rise of traffic. Sometimes, scammers using multiple combination of common name at popular domain name to send scam[3].

The risks of not filtering spam are the constant flood of spam networks clogs and corresponding impacts on user inboxes, but also downgrade valuable resources such as bandwidth and storage capacity, productivity loss and interfere with the expedient delivery of legitimate emails. Not only is spam frustrating for most email users, it strains the IT infrastructure of both software and hardware of an organizations and costs businesses to lose billions of dollars in their productivity[4]

Today, Spammers are exploring the advantages of electronic mail (email) .This is because of its efficiency, effectiveness and it is considered very cheap as they can send the same messages to many email users from addresses gotten by various means. For example, the use of automatic programs called bots such as web crawlers or spiders to scour the Web and Usenet newsgroups, collecting addresses, or buy email addresses in bulk from other companies at very low prices. Thanks to spoofing, spammers are now able to defraud innocent and greedy victims.

In order to address the various growing problem in spam , organization must analyze the tools available to determine how best to counter spam in its environment[6]. Tools, such as the corporate e-mail system, e-mail filtering gateways, contracted anti-spam services, and end-user training, provide an important arsenal for any organization.

2. Related Work

The investigation on the usage of the word “spam” being associated with unsolicited commercial emails is not entirely clear. The fact that SPAM was created by Hormel in 1937 as the world’s first canned meat that didn’t need to be refrigerated. It was originally named “Hormel Spiced Ham,” but was eventually changed to the catchier name, “SPAM.” Its connection to email is, according to Hormel and many other sources, due to a sketch on the British comedy TV show, Monty Python’s Flying Circus. In the skit, a group of Vikings sing “SPAM, SPAM, SPAM” repeatedly, drowning out all other conversation in the restaurant.

An in-depth research into the history of spam on the internet was carried out by Brad Templeton, founder of ClariNet Communication Corp. According to him, the first email spam was from 1978, and was sent out to all users on ARPANET (several hundred users). It was an ad for a presentation by Digital Equipment Corp. Templeton notes that the origin of spam as we know started on Usenet and migrated to email.

Fabício B [9] used content filtering techniques whereby content are blocked or allowed based on analysis of its content rather than its source or other criteria. However there was no a clear security model standard designed to limit the extent of security incidents such as worms which could potentially overload the Internet causing a global denial of service.

Developing intelligent and sophisticated content filtering technology with standards and cooperation among ISPs may be the solution. Natarajan 2010[9] provide a third party large-scale blacklist to decide which email is spam. A blacklist is a list of traits that spam emails have, and if the email to be tested contains any of those traits, it is marked as spam. It is possible to organize a blacklist based on "From:" fields, originating IP addresses, the subject or body of the message, or any other part of the message that makes sense. A small-scale blacklist works fine if the user gets spam from one particular address. He was unable to provide a solution on a larger scale, where the user does not have any control over the blacklist, there must be a mechanism in place for dealing with accidental blacklisting of other users[10].

The report by O' Brien J and Chiarella J (2003)[11] state that it is obvious problem that it is impossible to predict who is going to send email, and anyone previously unknown to the user will be filtered out. One way to avoid this problem is to read through the filtered email regularly but there is no point in filtering if the user must view all of the email anyways.

Androutsopoulos, I[12] in his work define how Bayesian is different from others because of its learning. To decide that incoming mail is spam or not, the filter needs to know about the mail that user receives. Spam is kept in separate table and that probabilities can be calculated. In this case, the user must manually indicate whether that email is spam or not. To train the filter there should be an intelligent mechanism to investigate the required trained word. Greylisting is the technique to temporarily reject messages from unknown sender mail servers as reported in [13].

In a related review Clark et al. [14] presented automated E-mail systems that were able to fill up the incoming E-mail messages into folders and anti-spam using neural network based system. The investigations from the study reveal that the technique is more accurate than several other techniques. The proposed technique mainly deals with clustering or grouping of mails into appropriate folders, rather than e-mail filtering. Wu (2009)[15] used a hybrid method of rule-based processing and back-propagation neural network for spam filtering. A rule-based process is first employed to identify and digitize the spamming behaviors observed from the headers and syslogs of e-mails. Then they utilize the spamming behaviors as features for describing e-mails. This information is then used to train the BPNN. The system produced very low false positive and negative rates. Meizhen et al. (2009)[16] proposed a model for spam behavior recognition based on fuzzy decision tree (FDT). This model can efficiently detect and analyze spammers' behavior patterns, and classify e-mails automatically. They concluded that since absolutely clear attributes does not always exist in the real world, the attribute subordinating degree is more

natural and reasonable to describe the characteristics of behavior. Fuzzy decision tree is more adaptive than Crisp decision tree.

In the aforementioned related research work, spam filtering methods is devised to work on the receiving end. Merely detecting a user sending out email after email and terminating their access would probably be sufficient to block spammers. The problem does not lie in detecting the spam. The problem is that some ISPs are willing to let spammers use their service to send out thousands of emails.

The report in this paper adopts the principle of quantitative and qualitative. The principle of the quantitative technique is asking as much respondents as possible to get adequate results of the research while quantitative is the method of data collection chosen in concordance with the explained methods.

3. Methodology

This section presents a complete proposed system design, deduced system requirement and implementation. The report in this paper adopts the principle of quantitative and qualitative. The principle of the quantitative technique is asking as much respondents as possible to get adequate results of the research while quantitative is the method of data collection chosen in concordance with the explained methods. Most email service providers such as yahoo mail, Gmail, implements a spam filtering application to detect spam messages. For incoming mail, the spam filtering application will check the mail and determine whether to place it in the spam box or inbox of the intended recipients. The traditional existing system for incoming E-mail is depicted in Figure 1 below.

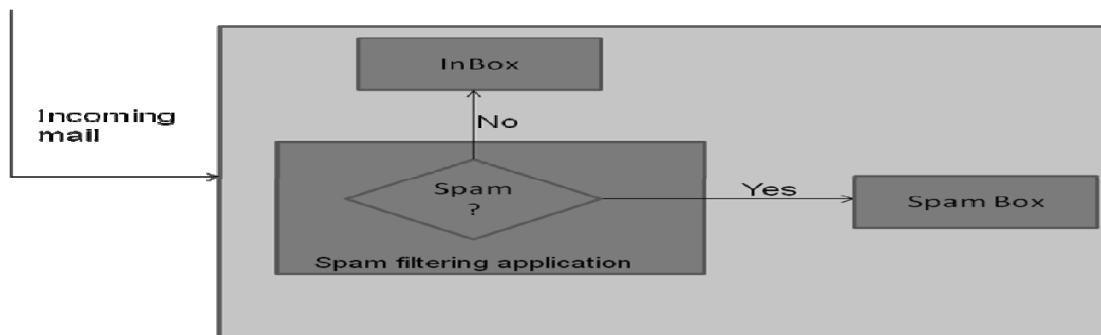


Fig 1: Existing System for Incoming Mail

For outgoing mail, when mail is sent by a user, it doesn't go through any form of spam checks from the system. Instead, it is sent out of the system as shown in Fig2 below.

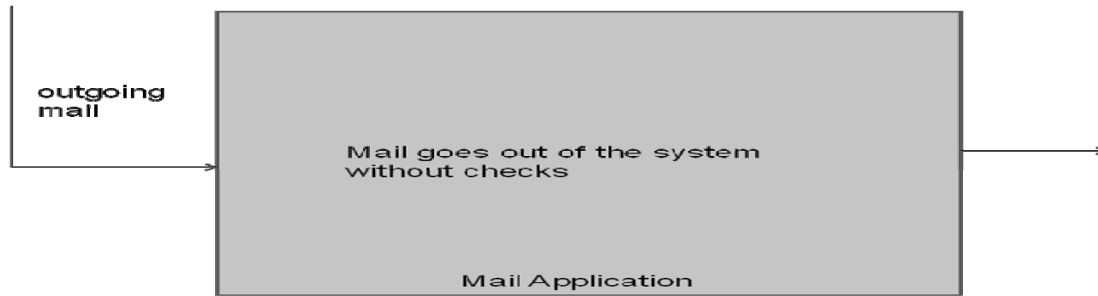


Fig 2: Existing System for outgoing Mail

The current system is very faulty because it only allows an administrator to deactivate the system users. It does not block or disallow system users from sending the SCAM message. Thus, it is imperative check through the messages and determine whether it is spam or not and then take the necessary actions.

3.1 Architecture of the Proposed Spam Filtering System Design

The proposed system Architecture is based on Bayesian techniques that uses mathematical formulae to analyze the content of a message, learning from the user which is a valid message and which is spam. Bayesian spam filtering is the process of using Bayesian statistical methods to classify documents into categories. Using well known mathematics, it is possible to generate a “spam indicate probability” for each word. Bayesian is different from others because of its learning process. To decide that incoming mail is spam or not, the filter needs to know about the mail that user receives. Spam is kept in separate table and that probabilities can be calculated. Bayesian rule using this probability: for example, most email users encounter the word ‘Viagra’ in spam email, but rarely want it in other email. The filter doesn’t know these probabilities in advance and must be trained first so it can build them up. A Bayesian spam filter relies on two things to work effectively: how well the Bayesian analysis formula has been implemented and how good a sample of data it has to work with. According to Wikipedia (2011), Bayesian spam filtering is the process of using Bayesian statistical methods to classify documents into categories. Using well known mathematics, it is possible to generate a “spam indicate probability” for each word.

Using Bayes’ theorem, one can conclude according to equation [j] that:

$$P(\text{spam} | \text{words}) = P(\text{words} | \text{spam})P(\text{spam}) / P(\text{words}) \dots\dots\dots \text{Eq. j}$$

Where $P(\text{spam} | \text{words})$ is the probability of spam where there is word

$P(\text{words} | \text{spam})$ is the probability of word where there is spam

$P(\text{spam})$ is the probability of spam

$P(\text{word})$ is the probability of word

3.12 Bayesian Statistical Scam Filter of the Proposed Design

In probability theory and statistics, Bayes' theorem (alternatively Bayes' law or Bayes' rule) is a method of incorporating new knowledge to update the value of the probability of occurrence of an event. To that end the theorem gives the relationship between the updated probability $P(A | B)$, the conditional probability of A given the new knowledge B , and the probabilities of A and B , $P(A)$ and $P(B)$, and the conditional probability of B given A , $P(B | A)$. In its most common form, Bayes' theorem is:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}.$$

Based on the theoretical background of Bayesian theory and provided spam (scam or non scam) is obtained, equation [K] is derived.

$$P(\text{scam} | \text{non scam}) = P(\text{non scam} | \text{scam})P(\text{scam}) / P(\text{non scam}) \dots \dots \dots \text{Eq. K}$$

Where

$P(\text{scam} | \text{non scam})$ is the probability of scam where there is non scam

$P(\text{non scam} | \text{scam})$ is the probability of non scam where there is scam

$P(\text{scam})$ is the probability of scam

$P(\text{non scam})$ is the probability of non scam

3.1 Proposed System Design of The Intelligent Spam Filtering

The analysis of the traditional existing system deduced that most email service providers such as yahoo mail, Gmail, implements a spam filtering application to detect spam messages but they have no scam filtering application to detect scam messages prioritization. For incoming mail, the spam filtering application will check the mail and determine whether to place it in the spam box or inbox, for messages intended for the inbox, our Bayesian statistical scam filtering application will determine whether to place it in the scam box or inbox of the intended recipients. For outgoing mail, when mail is sent by a user, it goes through the Bayesian spam filtering that will be implemented. If the mail is not scam, it goes out of the system and email service providers such as yahoo mail and Gmail checks whether the mail is spam or not with the spam filtering application which will determine if the mail is to be placed in spam box or inbox . If the mail is scam, it enters the scam net where only the administrator as access to. Sometimes, some mails that are not scam are mistakenly classified as one (false positives) if so, the administrator sends it out of the system. Also, mails that enter the scam net are used to train words in the knowledge base and new techniques employed by scammers are added to the knowledge base. The newly proposed system will attempt to SCAN through every message that is about to be sent to the

intended recipient and with some set of algorithms, determine whether the message is spam or not as shown in fig3.

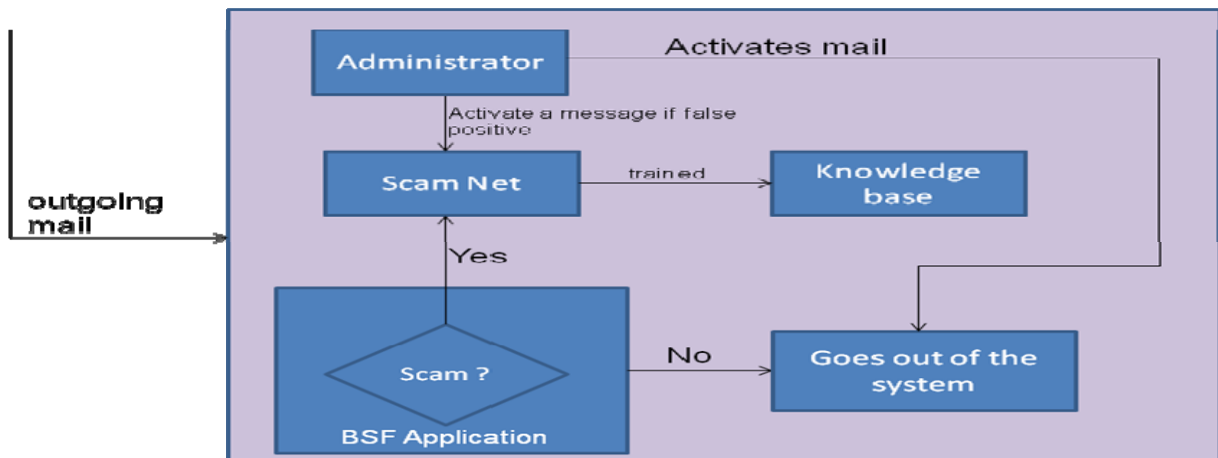


Fig 3: The proposed inbound spam filtering system

3.2 System Requirement of the Design

This section lists the specific and important requirement of the design including the various functions of the system and it contains snapshots of how the API is being used to detect SCAM.

Given a message **q**, that a user **u** is about to send, the system does the following to **q**

- ✓ Tokenises **q** to several words and places the token into an array
- ✓ Removes one and two letter words from the token array
- ✓ Remove neutral words from the token array
- ✓ Loop through the array to obtain the spamicity of each token
- ✓ Based on 4, the system applies the Bayesian theory to obtain the message spamicity.

Each of the process above will be explained in detail.

Process One: Tokenization

The message is divided into words using the separator **space**. So each of the token is then placed inside an array. Each of the token occupies a given location in the array. Tokenization process also involves removing duplicate entry of a token. A module in the SpamTrainer class handles this process and it is been implemented as follows:

```
$sp = new SpamTrainer($message,$messageType);  
$sp->tokenise();
```

Process Two: Remove One and two letter Words

This process removes every one and two letter token from the token array initially created from the first process. According to the Bayesian algorithm, such words are not relevant and should be discarded.

...

```
$sp->removeOneAndTwoletterWord ();
```

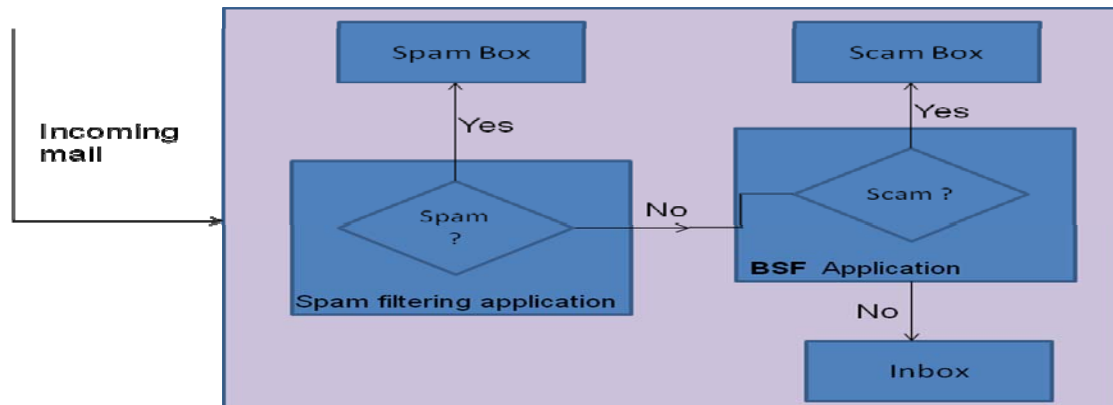


Figure 5: The Proposed outbound Spam Filtering System

Process Three: Remove Neutral Words

There are some words such as *some, which, I, for, are, all etc* are regarded as neutral words. These words are subsequently inserted in to the database by the SCAM monitor admin. The above process contacts the database for such words and removes them from the token previously created if they exist. Such words are not relevant in the determination of the message spamicity and hence they are discarded.

...

```
$sp->filterOutNeutralWords()
```

Process Four: Obtain Spamicity of Individual token

The main process and stages of spamicityof individual token for add up and return messages is broken as follows: Get the spamicity of each individual tokens and obtain the spamicity of the message as depicted in the proposed class diagram in fig 4 below.

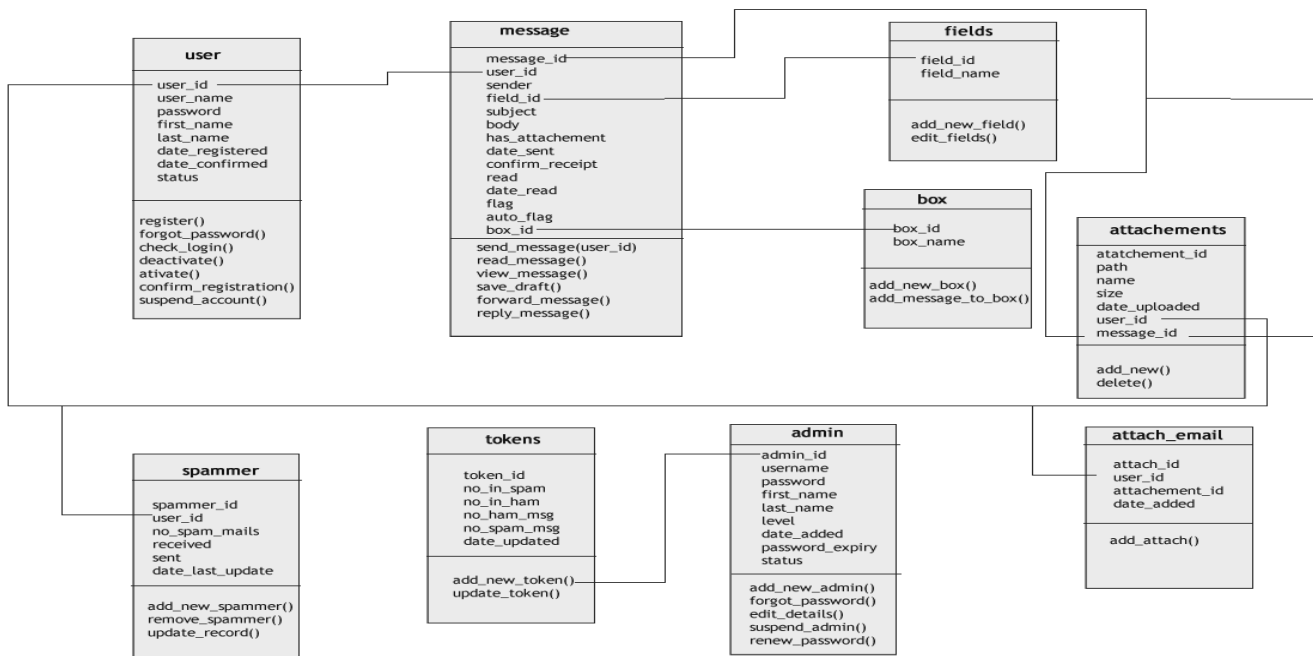


Fig 4: Proposed Class Diagram for the entire email system

The function **spamicity (\$word)** returns the spamicity of a given word. It does this by contacting the database for the word and applying the Bayesian conditional probability described in Chapter three to obtain the spamicity for the word. If the word is not in the database, a spamicity of **0.50** is awarded the word/token.

The function **getMessageSpamicity()** calls the function spamicity to execute for each of tokens inside the token array and then apply the Bayesian conditional probability described in section below to determine the message spamicity.

Field	Type
spammer_id	int(11)
user_id	int(11)
no_spam_mails	int(11)
received	int(11)
sent	int(11)
date_last_update	varchar(20)

Fig1: Relationship of database Table of Spammer

4.2 Choice for Threshold value

The threshold value set is 0.50. This is because for any word that does not exist in the database, the function spamicity returns a 0.50. Thus we can say that anything above 0.50 is definitely SCAM. The getMessageSpamicity() function returns the spamicity of the messages passed into the class. Thus

Given that for a message **q** and type null, we decide to obtain the spamicity, the following steps are obtained.

```
//tokenise the message
$this->spamTrainer->tokenise();

//filter out neutral words
$this->spamTrainer->filterOutNeutralWords();
//get spamicity of message
$spamicity = $this->spamTrainer->getMessageSpamicity();

If($spamicity > 0.50)
{ return "message is spam";}
```

5.1 Discussion of Results

The discussion of the proposed model is centered on our requirement for the design such as the GUI components containing a mysql server as the database make up the scam filtering system . API was used to develop the system so that any Internet or E-mail Service Provider can easily integrate the system with their existing system. We experiment with Yahoo mail and Google mail using the same ham and scam messages. The GUI results of yahoo mail and Google are presented in Figure 7 and Figure 8 while the figure 5 and 6 show the results of our proposed application with the system developed messages to determined the spamicity obtained based on 0.5 threshold. Based on the threshold, 0.5 set in scam filter, the scam filtering system recorded a True Classification Rate (sum of True Positive and True Negative) of 0.9 on the ten (10) scam messages tested. A sample of this is presented in figure 5 and 6 showing how our e-mail scam filtering system correctly classified messages. Message in figure 5 was classified as ham and placed in Inbox because its spamicity was 0.500 while message in figure 6 was classified spam because its spamicity of 0.930 exceeds 0.5.

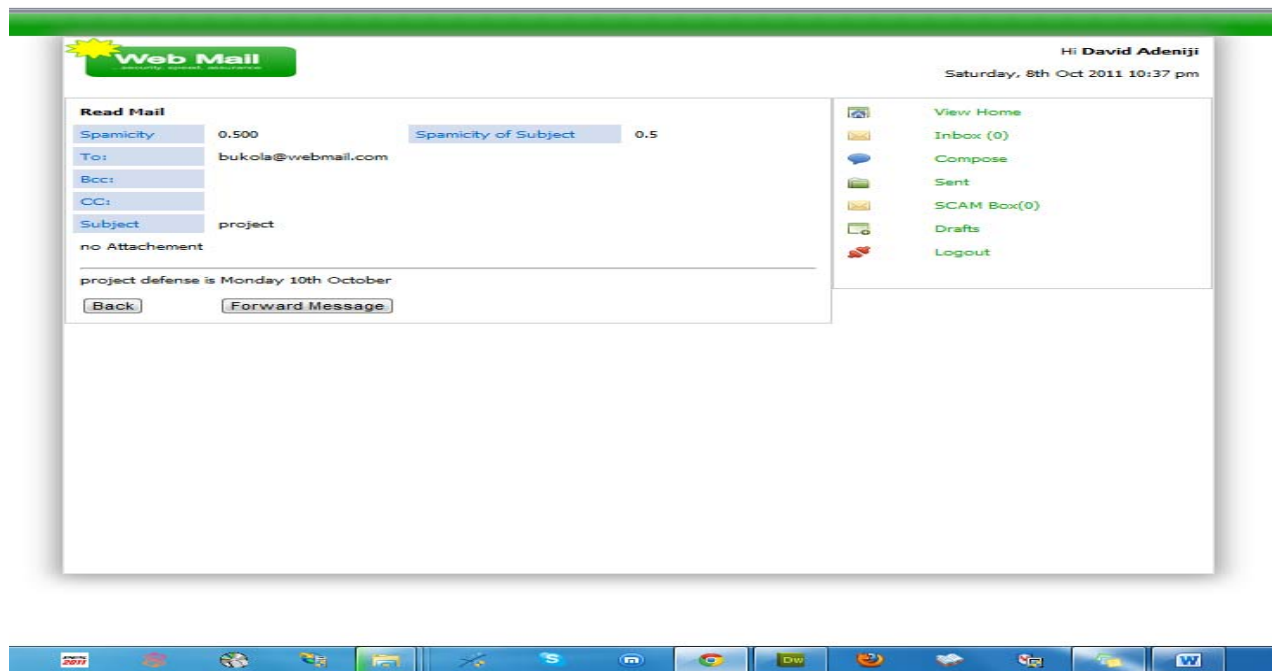


Fig5: Read mails from Inbox

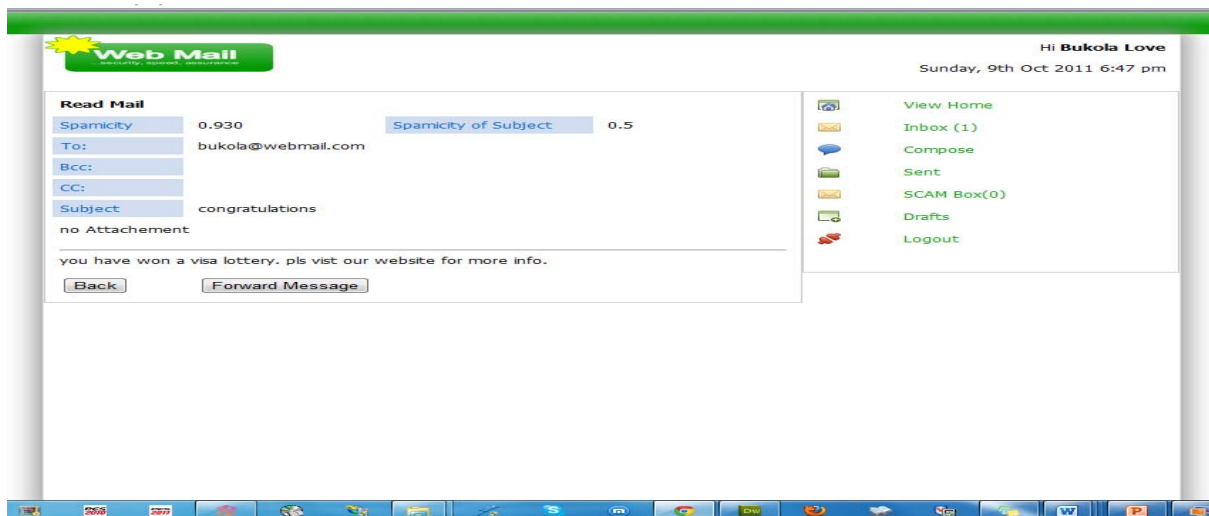


Fig6: Read mail from Scam Box

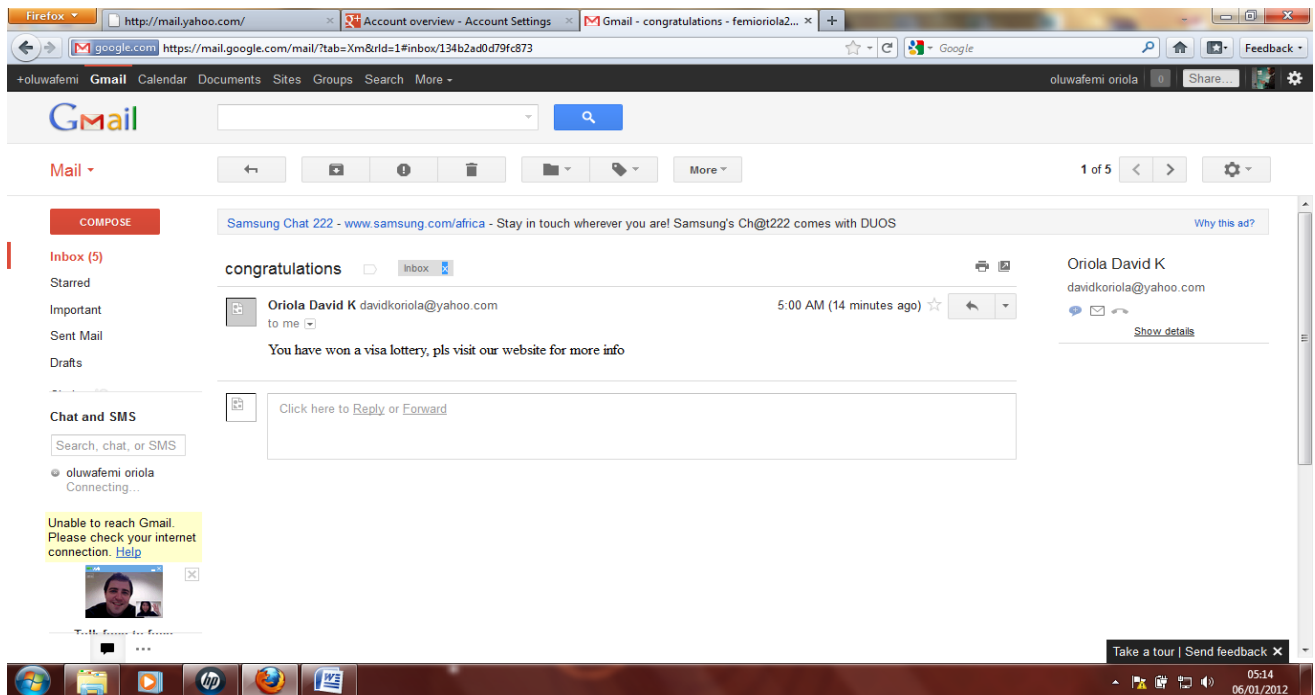


Fig7: Read Mail from Gmail Inbox

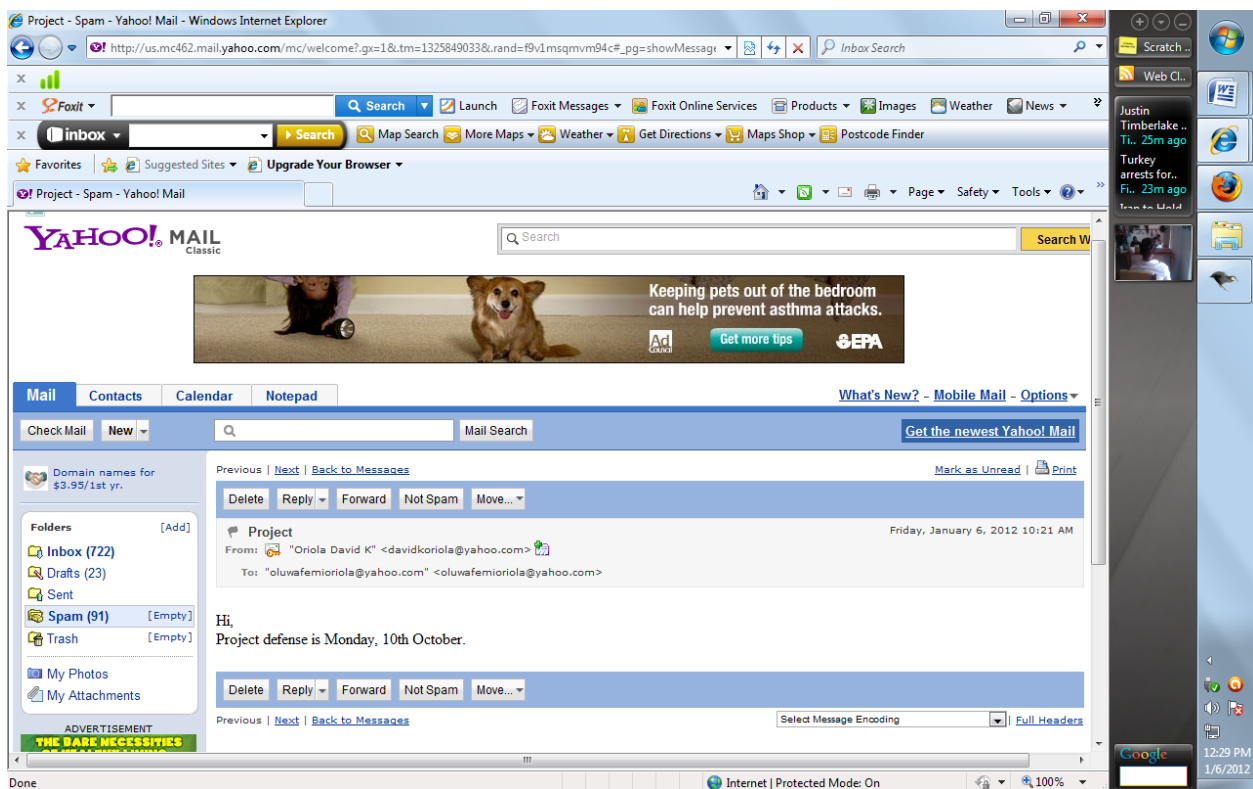


Fig8: Read Mail from Yahoo Spam Box

Table 2 shows the summary of the evaluation on 20 messages comprising of ten (10) hams and ten (10) scams.

Table 2: Summary of Evaluation of E-Mail Systems on Filtering of Scam

	Scam Filtering System	Yahoo Mail	Google Mail
Scam	9	6	7
Inbox	11	14	13

Yahoo mail Guards recorded a True Classification Rate of 0.6 with same set of messages given the same knowledge base. Google mail spam filter True Classification Rate was 0.7 with the same condition. The summary of the evaluation presented in table 2 shows that our scam filter outperforms Yahoo mail Spam Guard and Google Mail Spam. However, Google mail Spam Filter performs better than Yahoo mail Spam Guard.

CONCLUSION

The new intelligent system is designed to meet the local INTERNET providers' needs such as an automated view of activity logs of every action carried out by a user, deactivation and activation of clients, auto-train software with new words . The SCAM filter software is also designed to remove every form of flooding and illegal spoofing. Over time we have seen scammers sending messages in the name of other legitimate company there by misleading innocent recipients. The SCAM detector automatically blocks messages from service provider because such messages wouldn't have been sent on a local INTERNET service provider's web application.

Based on the conclusions and findings of this study it is noticed that the fight against SCAM messages on EMAIL web application programs is an interesting and growing area of research which could be further investigated to include a variety of functionalities. The scope of the work did not cover for BULK SMS messages. BULK SMS is very cheap and these spammers always try to take advantage of this to defraud innocent citizens. Work is being going on the topic but there are still some areas such as detecting image SCAMs which is still ongoing.

REFERENCES

1. Ahmad, B. B. I., 2007. "Spam Filtering Implementation Using Open Source Software" Accessed 8th September, 2011.
2. Chih-Chien, W., 2003. "Sender and Receiver Addresses as Cues for Anti-Spam Filtering," Journal of Research and Practice in Information Technology, 36(1), 3-7.
3. USIC3 - Internet Crime Complaint Centre Report (2006-2008). www.ic3.gov/media/annualreports.aspx. Accessed 8 September,.

4. Fabrício B., Tiago R., Virgílio A., Jussara A & Marcos G. (2009); DETECTING SPAMMERS AND CONTENT PROMOTERS IN ONLINE VIDEO SOCIAL NETWORKS; In ACM SIGIR Conference, Boston, MA, USA, July 2009.
5. Pantel, P., 1998. "Spamcop—a spam classification and organization program," Proceedings of AAAI-98 Workshop on Learning from Text Categorization. 1998.
6. Sakkis, G., Androutsopoulos, I., Paliouras, G., Karkaletsis, V., Spyropoulos, C. D., & Stamatopoulos, P., 2003. A memory-based approach to anti-Spam filtering for mailing lists. *Information Retrieval*. 6(1), 48–73.
7. Androutsopoulos, I., Paliouras, G., Karkaletsis, V., Sakkis, G., Spyropoulos, C., and Stamatopoulos, P. 2000c. Learning to filter spam e-mail: A comparison of a naive Bayesian and a memory-based approach. In Proceedings of the Workshop on Machine Learning and Textual Information Access, 4th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2000) (Lyon, France), H. Zaragoza, P. Gallinari, and M. Rajman, Eds. 1—13
- 8 Symantec Global Internet Security Threat Report Trends for 2010 Volume XVI, Published April 2010. <http://www.symantec.com/connect/2011>. Accessed 30 June, 2011
9. Natarajan Arulanand (2010): Payload Inspection Using Parallel Bloom Filter in Dual Core Processor; *Computer and Information Science: Vol. 3, No. 4; 2010*.
10. Rajkumar, B., Tianchi, M., Rei, S., Chris, S., and Willy, S., 2006, "Domain Specific Blacklists," Proceedings of the Fourth Australian Information Security Workshop (AISW-NetSec 2006).
- 10 Natarajan Arulanand (2010): Payload Inspection Using Parallel Bloom Filter in Dual Core Processor; *Computer and Information Science: Vol. 3, No. 4; 2010*.
11. O' Brien J and Chiarella J (2003): AN ANALYSIS OF SPAM FILTERS; Available at; <http://web.cs.wpi.edu/~claypool/mqp/spam/mqp.pdf> on 9/10/2011. 1
12. Androutsopoulos I, Koutsias J, Chandrinos KV, Paliouras G, Spyropoulos C (2000). An evaluation of naïve Bayesian anti-spam filtering. Proc. Of the workshop on machine learning in the new information age: 11th Europe conference on machine learning, pp. 9- 17.
13. Sender and Receiver Addresses as Cues for Anti-Spam Filtering. *Journal of Research and Practice in Information Technology*, 36(1), 3-7.
14. Clark, J.; Koprinska, I.; and Poon, J. (2003). A neural network based approach to automated e-mail classification. *Proceedings of Web Intelligence, Proceedings. IEEE/WIC International Conference*, 702- 705.
15. Wu CH (2009). Behavior-based spam detection using a hybrid method of rule-based techniques and neural network. *Expert Systems with Application*. Elsevier, pp. 4321-4330.
- 16 Meizhen W, Zhitang L, Sheng Z (2009). A Method for Spam Behavior Recognition Based on Fuzzy Decision Tree. *IEEE, Ninth International Conference on Computer and Information Technology*, pp. 236-241..

New Weather Forecasting Technique using ANFIS with Modified Levenberg-Marquardt Algorithm for Learning

I.Kadar Shereef

Head, Department of Computer Applications
Sree Saraswathi Thyagaraja College, Pollachi.
Coimbatore, Tamil Nadu, India.
kadarshereef@gmail.com

Dr. S. Santhosh Baboo

Reader, PG and Research department of Computer Science
Dwaraka Doss Goverdhan Doss Vaishnav College
Chennai, Tamil Nadu, India

Abstract---Temperature warnings are essential forecasts since they are utilized to guard life and property. Temperature forecasting is the kind of science and technology to approximate the temperature for a future time and for a given place. Temperature forecasts are performed by means of gathering quantitative data regarding the in progress state of the atmosphere. The author in this paper utilized a neural network-based technique for determining the temperature in future. The Neural Networks package consists of various kinds of training or learning techniques. One such technique is Adaptive Neuro Fuzzy Inference System (ANFIS) technique. The main advantage of the ANFIS technique is that it can reasonably estimated a large class of functions. This technique is more efficient than numerical differentiation. The simple meaning of this term is that the proposed technique has ability to confine the complex relationships among several factors that contribute to assured temperature. The proposed idea is tested using the real time dataset. In order to further improve the prediction accuracy, this paper uses Modified Levenberg-Marquardt (LM) Algorithm for Neural Network learning. In modified LM, the learning parameters are modified. The proposed algorithm has good convergence and also it reduces the amount of oscillation in learning procedure. The proposed technique is compare with the usage of ANFIS and the practical working of meteorological department. The experimental result shows that the proposed technique results in better accuracy of prediction when compared to the conventional technique of weather prediction.

Keywords--- Multi Layer Perception, Temperature Forecasting, Back propagation, Artificial Neural Network, Modified Levenberg-Marquardt Algorithm

1. INTRODUCTION

THE enormous computational is necessary to resolve the equations that represents the atmosphere, error concerned in measuring the initial conditions, and an imperfect understanding of atmospheric procedures because of chaotic nature [8, 20] of the atmosphere. This indicates that forecasts turn out to be less precise as the dissimilarity

in current time and the time for which the forecast is performed (the range of the forecast) increases. The use of ensembles and model helps narrow the error and pick the most likely outcome.

Various proves involved in temperature prediction are

- Data collection(atmospheric pressure, temperature, wind speed and direction, humidity, precipitation),
- Data assimilation and analysis
- Numerical weather prediction
- Model output post processing

A neural network [1] is a dominant data modeling technique that has the capability to capture and symbolize complex input /output relationships. The inspiration for the growth of neural network is obtained from the aspiration to realize an artificial system that could carry out intelligent works related to those carry out by the human brain. Neural network look like the human brain in the following manners:

- A neural network acquires knowledge through learning
- A neural network's knowledge is stored within interneuron connection strengths known as synaptic weights

The exact supremacy and merits of neural networks [12] occurs in the capability to symbolize both linear and non linear relationships straightforwardly from the data being modeled. Conventional linear models are simply insufficient when it approaches for true modeling data that consists of non linear features.

A neural network model is a formation that can be altered to result in a mapping from a provided set of data to characteristics of or relationships between the data. The model is modified, or trained, with the help of collection of data from a provided source as input, usually referred to as the training set. When the training phase completed successful, the neural network will be capacity to carry out

classification, estimation, prediction, or simulation on new data from the same or similar sources.

An Artificial Neural Network (ANN) [2, 4, 5] is a data processing model that is motivated by the manner biological nervous systems like the brain, process those data. The main constituent of this model is the new structure of the data processing system. It consists of a large number of extremely interrelated processing elements (neurons) functioning together in order to resolve particular problems. ANNs, like people, are trained by illustrations. An ANN is constructed for some application like pattern recognition or data classification, by means of a learning process. Learning in biological systems provides alterations to the synaptic relation that occurs among the neurons.

A back propagation network [9] contains at least three layers (multi layer perception):

- An input layer
- At least one intermediate hidden layer
- An output layer

In distinction to the Interactive Activation and Competition (IAC) Neural Networks and Hopfield Networks, relation weights in a back propagation network are single way. Normally, input units are linked in a feed-forward manner with input units completely linked to units in the hidden layer and hidden units completely linked to units in the output layer. An input pattern is transmitted forward to the output units by means of the intervening input-to-hidden and hidden-to-output weights when a Back Propagation network is cycled.

As the algorithm's name provides a meaning, the errors (and consequently the learning) propagate backwards from the output nodes towards the inner nodes. Therefore precisely it can be explained, back propagation is utilized to compute the gradient of the error of the network with regard to the network's adjustable weights. This gradient is forever utilized in a simple stochastic gradient descent technique to identify weights that reduces the error. Regularly the term back propagation is mentioned in a more common means in order to mention the complete process surrounding both the computation of the gradient and its utilization in stochastic gradient manner. Back propagation frequently permits fast convergence on acceptable local minima for error in the type of networks to which it is suited.

The projected Temperature Prediction System which utilizes BPN Neural Network and [13-16] modified LM algorithm [22] is evaluated with the help of the dataset from [17]. The results are contrasted with practical temperature prediction outcome [18, 19]. This system supports the meteorologist to forecast the expectation weather effortlessly and accurately.

The remainder section of this paper is organized as follows. Section 2 discusses various temperature predicting systems with various learning algorithms that were earlier proposed in literature. Section 3 explains the proposed work of developing An Efficient Temperature Prediction System using ANFIS with modified LM algorithm. Section 4 illustrates the results for experiments conducted on sample

dataset in evaluating the performance of the proposed system. Section 5 concludes the paper with fewer discussions.

2. RELATED WORK

Several works were performed related to the temperature prediction system and BPN network conventionally. Some of the works summarized below.

Y. Radhika *et al.*, [3] presents an application of Support Vector Machines (SVMs) for weather prediction. Time series data of every day maximum temperature at place is considered to forecast the maximum temperature of the successive day at that place according to the every day maximum temperatures for a period of earlier n days referred to as organize of the input. Significance of the system is practical for different spans of 2 to 10 days with the help of optimal values of the SVM kernel.

Mohsen Hayati *et al.*, [5] studied about Artificial Neural Network based on MLP was trained and tested using ten years (1996-2006) meteorological data. The outcome suggests that MLP network has the lesser prediction error and can be recognized as a better technique to model the short-term temperature forecasting [STTF] systems. Brian A. Smith *et al.*, [6] aims at creating a ANN models with lesser average prediction error by means of enhancing the number of distinct observations utilized in training, adding together extra input expressions that explain the date of an observation, raising the duration of prior weather data considered in all observation, and reexamining the number of hidden nodes utilized in the network. Models were generated to predict air temperature at hourly intervals from one to 12 hours before it happens. The entire ANN model, containing a network architecture and set of associated parameters, was calculated by instantiating and training 30 networks and computing the mean absolute error (MAE) of the resulting networks for few set of input patterns.

Arvind Sharma *et al.*, [7] briefly provided the way of the various connectionist models could be created with the help of various learning techniques and then examines whether they can afford the necessary level of performance, that are adequately good and robust so as to afford a reliable prediction model for stock market indices.

Mike O'Neill [11] considers two major practical concerns: the relationship among the amounts of training data and error rate (equivalent to the attempt to collect training data to create a model with provided maximum error rate) and the transferability of models' expertise among various datasets (equivalent to the helpfulness for common handwritten digit recognition). Henry A. Rowley reduces the complicated work of manually choosing nonface training illustrations, that must be preferred to period the entire space of nonface images. Simple heuristics, like utilizing the detail that faces infrequently overlie in images, can additionally enhance the accuracy. Contrasting with more than a few other state-of-the-art face

detection techniques, it can be observed that the proposed system has better performance by means of detection and false-positive rates.

3. ANN APPROACH

A. Phases in Back propagation Technique

The back propagation [10] learning technique can be separated into two phases:

- Propagation
- Weight Update

Phase 1: Propagation

Each propagation includes the following process:

1. Forward propagation of a training pattern's input is provided by means of neural network for the purpose of producing the propagation's output activations.
2. Back propagation of the output activations propagation by means of the neural network with the help of training pattern's target for the purpose creating the deltas of every output and hidden neurons.

Phase 2: Weight Update

For each weight-synapse:

1. Multiply its input activation and output delta to obtain the gradient of the weight.
2. Bring the weight in the direction of the gradient by means of adding a proportion of it from the weight.

This proportion bangs on the speed and quality of learning; it is known as learning rate. The indication of the gradient of a weight assigns where the error is increasing; this is main reason for the weight to be updated in the reverse direction.

The phase 1 and phase 2 is continual until the performance of the network is acceptable.

B. Modes of Learning

There are fundamentally two kinds of learning to select from, one is on-line learning and the other is batch learning. Every propagation is followed straight away by means of a weight update in online learning [21]. In batch learning, much propagation happens before weight updating carried out. Batch learning requires extra memory capacity, but on-line learning needs more updates.

C. Basic ANFIS Architecture

Jang [JAN93] proposed ANFIS derived from Adaptive Network Based Fuzzy Inference Engine. This technique was intended to facilitate if-then rules and

membership function to be built depending on the historical data of the metrics. It also comprise the adaptive nature for automatic tuning purposes.

Figure 6.1 shows the basic architecture of ANFIS with two inputs and one output. ANFIS is a multilayer feed-forward network in which each node will execute a specific function on the incoming input signals. Each node will adapt and trained by altering its parameters and / or formulas. [JAN93] Proposed that the functions of the nodes are group into 5 different layers.

The back propagation equations are provided below. The equation (1) represents the way to compute the partial derivative of the error E^P regarding the activation value y^j at the n -th layer.

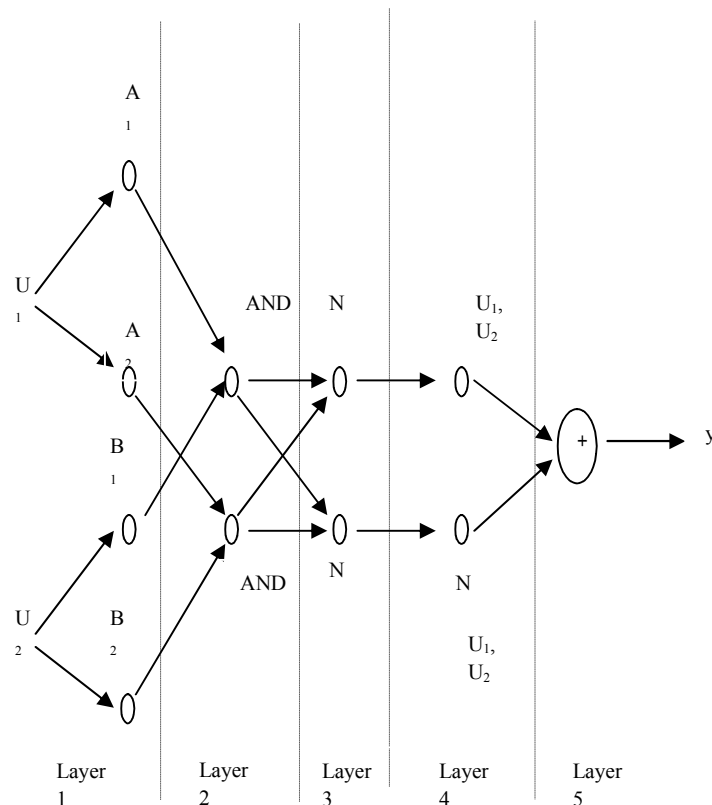


Figure-6.1: Basic Architecture of ANFIS

Initialize the procedure by calculating the partial derivative of the error because of a single input image pattern regarding the outputs of the neurons on the last layer. The error occurred because of the single pattern is computed as below:

$$E_n^P = \frac{1}{2} \sum \left[(x_n^I - T_n^I)^2 \right] \quad (1)$$

Where:

E_n^p represents the error because of a single pattern P at the last layer n ;

T_n^i represents the target output at the last layer (i.e., the desired output at the last layer) and x_n^i is the actual value of the output at the last layer.

Provided equation (1), then taking the partial derivative results in:

$$\frac{\partial E_n^p}{\partial x_n^i} = x_n^i - T_n^i \quad (2)$$

Equation (2) gives us a starting value for the back propagation process. The numeric values are used for the quantities on the right side of equation (2) in order to calculate numeric values for the derivative. Using the numeric values of the derivative, the numeric values for the changes in the weights are calculated, by applying the following two equations (3) and then (4):

$$\frac{\partial E_n^p}{\partial y_n^i} = \frac{G(x_n^i)(\partial E_n^p)}{\partial x_n^i} \quad (3)$$

where $G(x_n^i)$ is the derivative of the activation function.

$$\frac{\partial E_n^p}{\partial w_{n-1}^{ij}} = x_{n-1}^j \frac{\partial E_n^p}{\partial y_n^i} \quad (4)$$

Subsequently, using equation (2) once more and also equation (3), the error for the previous layer is computed, with the help of following equation:

$$\frac{\partial E_{n-1}^p}{\partial x_{n-1}^k} = \sum_i w_{n-1}^{ik} \frac{\partial E_n^p}{\partial y_n^i} \quad (5)$$

The values resulted from equation (5) are utilized as starting values for the computation on the directly preceding layer. This is the single most significant point in understanding back propagation. Otherwise it can be said that, it is taken the numeric values resulted from equation (5), and utilize them in a repetition of equations (3), (4) and (5) for the instantly preceding layer.

Simultaneously, the values resulted from equation (4) suggests the range to alter the weights in the current layer n , that was the entire reason of this gigantic exercise. Especially, the value of each weight is updated based on the following equation:

$$(w_n^{ij})_{new} = (w_n^{ij})_{old} - \eta \alpha_i \left(\frac{\partial E_n^p}{\partial w_n^{ij}} \right) \quad (6)$$

where η represents the learning rate, characteristically a small number like 0.0005 and will be decreased steadily during training.

The learning can be enhanced to improve the performance of prediction system. For this reason, this paper uses Modified Levenberg-Marquardt algorithm for learning phase of ANFIS.

ANFIS Algorithm:

Step 1: Layer 1: Here, the membership function are defined hypothetically and bell-shaped is generally selection, represented in equation below:

$$u_{A_i}(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right)^2 \right]^{\frac{1}{n_i}}}$$

When the values alter, the bell-shaped function will also change consequently. In this layer, the parameters present in the process are called as the premise parameters.

$$a_i = \frac{a_1}{a_1 + a_2}$$

Step 2: Layer 2: In this layer, each output of the node defined the firing strength of the rules in the fuzzy inference engine.

Step 3: Layer 3: This layer computes the ratio of the i^{th} rule's firing strength, as shown in equation (6.2). The results are the normalized firing strength.

Step 4: Layer 4: The parameters of the nodes in this layer are called the consequent parameters. The nodes in this layer adapts with an output node.

Step 5: Layer 5: Nodes in this layer are fixed and sums all incoming signals from the previous layers.

C. Modified Levenberg-Marquardt algorithm

A Modified Levenberg-Marquardt algorithm is used for training the neural network.

The learning algorithm used for this proposed approach is Modified Levenberg-Marquardt algorithm. This algorithm is clearly discussed in the chapter 4.

In this section, the way Modified Levenberg-Marquardt algorithm employed for updating the ANFIS parameters is explained. The ANFIS has two types of parameters which need training, the antecedent part parameters and the conclusion part parameters. The membership functions are assumed Gaussian as in the below equation:

$$\mu_{A_i}(x) = \frac{1}{1 + \left\{ \left(\frac{x - c_i}{a_i} \right)^2 \right\}^{b_i}} \quad (7)$$

And their parameters are $\{a_i, b_i, c_i\}$, where a_i is the variance of membership functions and c_i is the center of MFs. Also b_i is a trainable parameter. The parameters of conclusion part are trained and here are represented with $\{p_i, q_i, r_i\}$. There are 3 sets of trainable parameters in antecedent part $\{a_i, b_i, c_i\}$, each of these parameters has N genes. Where, N represents the number of Membership Functions. The conclusion parts parameters $\{p_i, q_i, r_i\}$ also are trained during optimization algorithm.

Parameters are initialized randomly in first step and then are being updated using Modified Levenberg-Marquardt algorithms. In each iteration, one of the parameters set are being updated. i.e. in first iteration for example a_i are updated then in second iteration b_i are updated and then after updating all parameters again the first parameter update is considered and so on.

ANFIS with Modified Levenberg-Marquardt Algorithm:

Step 1: The training parameters of the ANFIS are updated according to Modified Levenberg-Marquardt algorithm which is given in the following steps.

Step 2: The weight factor is updated using the performance index $F(w) = e^T$ which is obtained using the Newton method. The update weight factor is given by the below equation

$$a_{i+1} = a_i - A_i^{-1} \cdot g_i \quad (8)$$

Step 3: The gradient is obtained using the following equation with the Jacobian matrix $J(w)$

$$\nabla F(x) = 2J^T e(w) \quad (9)$$

Step 4: The Hessian Matrix is obtained by the following equation

$$\nabla^2 F(w) = 2J^T(W) \cdot J(W) + S(W) \quad (10)$$

Step 5: IF trial $S(w)$ in step 4 is very small, THEN Hessian Matrix is updated as

$$\nabla^2 F(w) \cong 2J^T(w)J(w) \quad (11)$$

Step 6: Using results of Step 5 and Step 2, Gauss-Newton method is obtained using the following equation

$$\begin{aligned} a_{i+1} &= a_i - [2J^T(a_i) \cdot J(a_i)]^{-1} 2J^T(a_i) e(a_i) \\ &\cong a_i - [J^T(a_i) \cdot J(a_i)]^{-1} J^T(a_i) e(a_i) \end{aligned} \quad (12)$$

Step 7: Gauss-Newton method is the matrix $H = J^T J$ is not invertible

Step 8: Then Hessian Matrix is modified for using the following equation

$$G = H + \mu I \quad (13)$$

Step 9: If the Eigen values and Eigen vectors of H are $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ and $\{z_1, z_2, \dots, z_n\}$ THEN

$$Gz_i = [H + \mu I]z_i \text{ and } Gz_i = (\lambda_i + \mu)z_i \quad (14)$$

Eigenvectors of G are the same as the eigenvectors of H, and the eigen values of G are $(\lambda_i + \mu)$.

Step 10: Matrix G is positive definite by increasing μ until $(\lambda_i + \mu) > 0$ for all i therefore the matrix will be invertible.

Step 11: In the standard LM method, μ is a constant number. In this modified LM, μ is modified as:

$$\mu = 0.01 e^T e \quad (15)$$

Thus e is a $k \times 1$ matrix therefore $e^T e$ is a 1×1 therefore $[J^T J + \mu I]$ is invertible

Step 12: After updating a_i , then b_i and c_i are updated similarly.

As known, learning parameter, μ is illustrator of steps of actual output movement to desired output. In the standard LM method, μ is a constant number. This research work LM method is modified using μ as

$$\mu = 0.01 e^T e \quad (16)$$

Where e is a $k \times 1$ matrix therefore $e^T e$ is a 1×1 therefore $[J^T J + \mu I]$ is invertible.

Therefore, if actual output is far than desired output or similarly, errors are large so, it converges to desired output with large steps. Likewise, when measurement of error is small then, actual output approaches to desired output with soft steps. Therefore error oscillation reduces greatly. Thus, Modified Levenberg-Marquardt algorithm is used for ANFIS learning which enhances the performance of the prediction technique.

4. EXPERIMENTATION AND RESULT

To experiment the proposed system a Madras Minambak, India (VOMM)[17] contains the real time observation of the weather for a particular period of time. For this experiment, an observation of 2010 year is taken. The dataset contains many attributes such as Temperature, Dew Point, Relative Humidity (RH), Wind Direction (DIR), Wind Speed (SPD) and Visibility (VIS).

4.1 Experimental Setup

TABLE 4.1

ANFIS with Modified LM	Seasons			
	Winter	Pre-Monsoon	South-West Monsoon	North-East Monsoon
Number of Hidden Neuron	6	6	6	6
Number of Epochs	150	150	150	150
Activation Function Used in Hidden Layer	Tan-sig	Tan-sig	Tan-sig	Tan-sig
Activation Function Used Output Layer	pure linear	pure linear	pure linear	pure linear

The experimental set up for this paper considers four seasonal variations. The available weather data were split into four seasons such as Winter (January-February), Pre-Monsoon (March-May), South-West Monsoon (June-September) and North-East Monsoon (October-December). This data is obtained from Indian Meteorological Department (IMD) [23]. In this experimental process, the missing values are obtained by the k-Nearest Neighbor algorithm.

Table 4.1 shows the various variables and parameters used for the ANFIS with Modified LM approach. The number of hidden neurons used in the present experimental observation is 6. Moreover, the number of iterations (epochs) taken is 150. The activation function used in Hidden and Output layer is Tan-sig and pure linear respectively for all the seasons considered.

Ten random days in each season are selected as unseen days. For Winter season, the unseen days chosen are 1/1/10, 2/1/10, 4/1/10, 18/1/10, 16/2/10, 20/2/10, 21/2/10, 23/2/10, 25/2/10 and 28/2/10. For Pre-Monsoon season, the unseen days chosen are 5/3/10, 8/3/10, 14/3/10, 27/3/10, 5/4/10, 10/4/10, 15/4/10, 18/5/10, 28/5/10 and 29/5/10.

For South-West Monsoon, the unseen days chosen are 6/6/10, 23/6/10, 29/6/10, 7/7/10, 19/7/10, 1/8/10, 20/8/10, 28/8/10, 2/9/10 and 27/9/10. For North-East Monsoon, the unseen days chosen are 1/10/10, 8/10/10, 28/10/10, 2/11/10, 15/11/10, 23/11/10, 29/11/10, 3/12/10, 14/12/10 and 25/12/10.

4.2 Performance Parameters

The performance of the proposed approaches are evaluated using the following parameters like

- Mean Squared Error (MSE)
- Minimum and Maximum Error and
- Prediction Accuracy

Mean Squared Error (MSE)

Table 4.2 shows the Mean Squared Error (MSE) comparison of the proposed approach and the existing approaches. The comparison is obtained for four seasons namely Winter, Pre-Monsoon, South-West Monsoon and North-East Monsoon.

Table 4.2
Mean Squared Error Comparison

Seasons	Mean Squared Error (Iterations =150)			
	BPN with Modified LM	Hybrid SOFM-MLP with Modified LM	ANFIS	ANFIS with Modified LM
Winter	0.083	0.067	0.017	0.0055
Pre-Monsoon	0.071	0.012	0.010	0.0034
South-West Monsoon	0.063	0.035	0.013	0.0046
North- East Monsoon	0.098	0.084	0.019	0.0065

For the South-West Monsoon season, the MSE obtained for the proposed BPN with LM approach is 0.063 which is very less than the MSE obtained by the existing approaches like BPN with LM and BPN with Linear Learning. South-West Monsoon season has the least MSE value.

The minimum and maximum error taken for four seasons are obtained and tabulated below table 4.3 , 4.4 and shows the minimum and maximum error comparison of the ANFIS approaches with various learning techniques.

Table 4.3: The Minimum error Comparison for four seasons

Seasons	Minimum Error			
	BPN with Modified LM	Hybrid SOFM-MLP with Modified LM	ANFIS	ANFIS with Modified LM
Winter	0.0093	0.0073	0.009	0.005
Pre-Monsoon	0.0089	0.0030	0.007	0.003
South-West Monsoon	0.0081	0.0052	0.008	0.004
North- East Monsoon	0.0097	0.0081	0.009	0.006

The minimum error obtained by the existing approaches such as BPN with Linear Learning and BPN with LM is higher when compared to the proposed BPN with Modified LM approach for all the seasons.

Table 4.4: The Maximum error Comparison for four seasons

Seasons	Maximum Error			
	BPN with Modified LM	Hybrid SOFM-MLP with Modified LM	ANFIS	ANFIS with Modified LM
Winter	0.6012	0.4220	0.2820	0.1230
Pre-Monsoon	0.5712	0.4002	0.2575	0.1053
South-West Monsoon	0.5392	0.4115	0.2725	0.1102
North- East Monsoon	0.6315	0.4352	0.2963	0.1334

Prediction Accuracy

Prediction accuracy for the proposed approaches for each season is tabulated in Table 4.5.

Table 4.5 Comparison of the Prediction Accuracy for Various Seasons

Seasons	Prediction Accuracy (%)			
	BPN with Modified LM	Hybrid SOFM-MLP with Modified LM	ANFIS	ANFIS with Modified LM
Winter	93.89	95.74	96.40	97.43
Pre-Monsoon	94.28	96.61	96.91	98.82
South-West Monsoon	94.87	96.10	96.55	98.19
North- East Monsoon	93.39	95.31	96.12	97.11

5. CONCLUSION

In this paper, ANFIS is used for predicting the temperature based on the training set provided to the neural network. Through the implementation of this system, it is illustrated, how an intelligent system can be efficiently integrated with a neural network prediction model to predict the temperature. This algorithm improves convergence and damps the oscillations. This method proves to be a simplified conjugate gradient method. When incorporated into the software tool the performance of the back propagation neural network was satisfactory as there were not substantial number of errors in categorizing. ANFIS approach for temperature forecasting is capable of yielding good results and can be considered as an alternative to traditional meteorological approaches. This paper uses Modified Levenberg-Marquardt Algorithm for Learning. This approach is able to determine the non-linear relationship that exists between the historical data (temperature, wind speed, humidity, etc.,) supplied to the system during the training phase and on that basis, make a prediction of what the temperature would be in future. The proposed approach is evaluated on Madras Minambak, India (VOMM) dataset. The performance of the proposed approach is evaluated based on the parameters like Mean Squared Error, Minimum and Maximum Error and Prediction Accuracy. The results are obtained and the values are tabulated for the data set. The performance of the proposed approach outperforms the existing three approaches based on the results obtained.

REFERENCES

- [1] Xinghuo Yu, M. Onder Efe, and Okyay Kaynak, "A General Back propagation Algorithm for Feedforward Neural Networks Learning,"
- [2] R. Rojas: Neural Networks, Springer-Verlag, Berlin, 1996.
- [3] Y.Radhika and M.Shashi, "Atmospheric Temperature Prediction using Support Vector Machines," International Journal of Computer Theory and Engineering, Vol. 1, No. 1, April 2009 1793-8201.
- [4] M. Ali Akcayol, Can Cinar, Artificial neural network based modeling of heated catalytic converter performance, Applied Thermal Engineering 25 (2005) 2341-2350.
- [5] Mohsen Hayati, and Zahra Mohebi, "Application of Artificial Neural Networks for Temperature Forecasting," World Academy of Science, Engineering and Technology 28 2007.
- [6] Brian A. Smith, Ronald W. McClendon, and Gerrit Hoogenboom, "Improving Air Temperature Prediction with Artificial Neural Networks" International Journal of Computational Intelligence 3;3 2007.
- [7] Arvind Sharma, Prof. Manish Manoria, "A Weather Forecasting System using concept of Soft Computing," pp.12-20 (2006)
- [8] Ajith Abraham¹, Ninan Sajith Philip², Baikunth Nath³, P. Saratchandran⁴, "Performance Analysis of Connectionist Paradigms for Modeling Chaotic Behavior of Stock Indices,"
- [9] Surajit Chattopadhyay, "Multilayered feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India," 2006
- [10] Raúl Rojas, "The back propagation algorithm of Neural Networks - A Systematic Introduction," chapter 7, ISBN 978-3540605058
- [11] Mike O'Neill, "Neural Network for Recognition of Handwritten

- Digits," Standard Reference Data Program National Institute of Standards and Technology.
- [12] Carpenter, G. and Grossberg, S. (1998) in Adaptive Resonance Theory (ART), The Handbook of Brain Theory and Neural Networks, (ed. M.A. Arbib), MIT Press, Cambridge, MA, (pp. 79–82).
 - [13] Ping Chang and Jeng-Shong Shih," The Application of Back Propagation Neural Network of Multi-channel Piezoelectric Quartz Crystal Sensor for Mixed Organic Vapours," Tamkang Journal of Science and Engineering, Vol. 5, No. 4, pp. 209-217 (2002).
 - [14] S. Anna Durai, and E. Anna Saro," Image Compression with Back-Propagation Neural Network using Cumulative Distribution Function," World Academy of Science, Engineering and Technology 17 2006.
 - [15] Mark Pethick, Michael Liddle, Paul Werstein, and Zhiyi Huang," Parallelization of a Backpropagation Neural Network on a Cluster Computer,"
 - [16] K.M. Neaupane, S.H. Achet," Use of backpropagation neural network for landslide monitoring," Engineering Geology 74 (2004) 213–226.
 - [17] <http://weather.uwyo.edu/cgi-bin/wyowx.fcgi?TYPE=sflist&DATE=current&HOUR=current&UNIT=A&STATION=VOMM>
 - [18] Grossberg, S.,"Adaptive Pattern Classification and Universal Recoding: Parallel Development and Coding of Neural Feature Detectors", Biological Cybernetics, 23, 121–134 (1976).
 - [19] Maurizio Bevilacqua, "Failure rate prediction with artificial neural networks," Journal of Quality in Maintenance Engineering Vol. 11 No. 3, 2005 pp. 279-294 Emerald Group Publishing Limited 1355-2511
 - [20] Chowdhury A and Mhasawade S V (1991), "Variations in Meteorological Floods during Summer Monsoon Over India", Mausam, 42, 2, pp. 167-170.
 - [21] Gowri T. M. and Reddy V.V.C. 2008. Load Forecasting by a Novel Technique using ANN. ARPN Journal of Engineering and Applied Sciences. 3(2): 19-25.
 - [22] Amir Abolfazl Suratgar, Mohammad Bagher Tavakoli and Abbas Hoseinabadi, "Modified Levenberg-Marquardt Method for Neural Networks Training", World Academy of Science, Engineering and Technology, Pp. 46-48, 2005.
 - [23] [http:// www.mausam.gov.in/WEBIMD/downloads/termglossory.pdf](http://www.mausam.gov.in/WEBIMD/downloads/termglossory.pdf).



Lt. Dr. S. Santhosh Baboo, aged forty two, has around Nineteen years of postgraduate teaching experience in Computer Science, which includes Six years of administrative experience. He is a member, board of studies, in several autonomous colleges, and designs the curriculum of undergraduate and postgraduate programmes. He is a consultant for starting new courses, setting up computer labs, and recruiting lecturers for many colleges. Equipped with a Masters degree in Computer Science and a Doctorate in Computer Science, he is a visiting faculty to IT companies. It is customary to see him at several national/international conferences and training programmes, both as a participant and as a resource person. He has been keenly involved in organizing training programmes for students and faculty members. His good rapport with the IT companies has been instrumental in on/off campus interviews, and has helped the post graduate students to get real time projects. He has also guided many such live projects. Lt. Dr. Santhosh Baboo has authored a commendable number of research papers

in international/national Conference/journals and also guides research scholars in Computer Science. Currently he is Senior Lecturer in the Postgraduate and Research department of Computer Science at Dwaraka Doss Goverdhan Doss Vaishnav College (accredited at 'A' grade by NAAC), one of the premier institutions in Chennai.



J. Kadar Shereef, done his Under-Graduation (B.Sc., Mathematics) in NGM College, Post-Graduation (MCA) in Trichy Jamal Mohamed College and Master of Philosophy Degree in Periyar University (distance education). He is currently pursuing his Ph.D., in Computer Science in Dravidian University, Kuppam, Andhra Pradesh. Also, he is working as a Lecturer, Department of BCA, Sree Saraswathi Thyagaraja College of Arts and Science, Pollachi. He is having more than two years of research experience and more than 6 years of teaching experience. His research interest includes Data mining, Climate Prediction, Neural Network and Soft Computing.

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore, Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGILIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Mrs Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Mr. P. Vasant, University Technology Petronas, Malaysia
Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Mr. Praveen Ranjan Srivastava, BITS PILANI, India
Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Mr. Tirthankar Gayen, IIT Kharagpur, India
Ms. Huei-Ru Tseng, National Chiao Tung University, Taiwan

Prof. Ning Xu, Wuhan University of Technology, China
Mr Mohammed Salem Binwahlan, Hadhramout University of Science and Technology, Yemen
& Universiti Teknologi Malaysia, Malaysia.
Dr. Aruna Ranganath, Bhoj Reddy Engineering College for Women, India
Mr. Hafeezullah Amin, Institute of Information Technology, KUST, Kohat, Pakistan
Prof. Syed S. Rizvi, University of Bridgeport, USA
Mr. Shahbaz Pervez Chattha, University of Engineering and Technology Taxila, Pakistan
Dr. Shishir Kumar, Jaypee University of Information Technology, Wakanaghat (HP), India
Mr. Shahid Mumtaz, Portugal Telecommunication, Instituto de Telecomunicações (IT) , Aveiro, Portugal
Mr. Rajesh K Shukla, Corporate Institute of Science & Technology Bhopal M P
Dr. Poonam Garg, Institute of Management Technology, India
Mr. S. Mehta, Inha University, Korea
Mr. Dilip Kumar S.M, University Visvesvaraya College of Engineering (UVCE), Bangalore University, Bangalore
Prof. Malik Sikander Hayat Khiyal, Fatima Jinnah Women University, Rawalpindi, Pakistan
Dr. Virendra Gomase , Department of Bioinformatics, Padmashree Dr. D.Y. Patil University
Dr. Irraivan Elamvazuthi, University Technology PETRONAS, Malaysia
Mr. Saqib Saeed, University of Siegen, Germany
Mr. Pavan Kumar Gorakavi, IPMA-USA [YC]
Dr. Ahmed Nabih Zaki Rashed, Menoufia University, Egypt
Prof. Shishir K. Shandilya, Rukmani Devi Institute of Science & Technology, India
Mrs.J.Komala Lakshmi, SNR Sons College, Computer Science, India
Mr. Muhammad Sohail, KUST, Pakistan
Dr. Manjaiah D.H, Mangalore University, India
Dr. S Santhosh Baboo, D.G.Vaishnav College, Chennai, India
Prof. Dr. Mokhtar Beldjehem, Sainte-Anne University, Halifax, NS, Canada
Dr. Deepak Laxmi Narasimha, Faculty of Computer Science and Information Technology, University of Malaya, Malaysia
Prof. Dr. Arunkumar Thangavelu, Vellore Institute Of Technology, India
Mr. M. Azath, Anna University, India
Mr. Md. Rabiul Islam, Rajshahi University of Engineering & Technology (RUET), Bangladesh
Mr. Aos Alaa Zaidan Ansaef, Multimedia University, Malaysia
Dr Suresh Jain, Professor (on leave), Institute of Engineering & Technology, Devi Ahilya University, Indore (MP) India,
Dr. Mohammed M. Kadhum, Universiti Utara Malaysia
Mr. Hanumanthappa. J. University of Mysore, India
Mr. Syed Ishtiaque Ahmed, Bangladesh University of Engineering and Technology (BUET)
Mr Akinola Solomon Olalekan, University of Ibadan, Ibadan, Nigeria
Mr. Santosh K. Pandey, Department of Information Technology, The Institute of Chartered Accountants of India
Dr. P. Vasant, Power Control Optimization, Malaysia
Dr. Petr Ivankov, Automatika - S, Russian Federation

Dr. Utkarsh Seetha, Data Infosys Limited, India
Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal
Dr. (Mrs) Padmavathi Ganapathi, Avinashilingam University for Women, Coimbatore
Assist. Prof. A. Neela madheswari, Anna university, India
Prof. Ganesan Ramachandra Rao, PSG College of Arts and Science, India
Mr. Kamanashis Biswas, Daffodil International University, Bangladesh
Dr. Atul Gonsai, Saurashtra University, Gujarat, India
Mr. Angkoon Phinyomark, Prince of Songkla University, Thailand
Mrs. G. Nalini Priya, Anna University, Chennai
Dr. P. Subashini, Avinashilingam University for Women, India
Assoc. Prof. Vijay Kumar Chakka, Dhirubhai Ambani IICT, Gandhinagar ,Gujarat
Mr Jitendra Agrawal, : Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal
Mr. Vishal Goyal, Department of Computer Science, Punjabi University, India
Dr. R. Baskaran, Department of Computer Science and Engineering, Anna University, Chennai
Assist. Prof, Kanwalvir Singh Dhindsa, B.B.S.B.Engg.College, Fatehgarh Sahib (Punjab), India
Dr. Jamal Ahmad Dargham, School of Engineering and Information Technology, Universiti Malaysia Sabah
Mr. Nitin Bhatia, DAV College, India
Dr. Dhavachelvan Ponnurangam, Pondicherry Central University, India
Dr. Mohd Faizal Abdollah, University of Technical Malaysia, Malaysia
Assist. Prof. Sonal Chawla, Panjab University, India
Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India
Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia
Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia
Professor Dr. Sabu M. Thampi, .B.S Institute of Technology for Women, Kerala University, India
Mr. Noor Muhammed Nayeem, Université Lumière Lyon 2, 69007 Lyon, France
Dr. Himanshu Aggarwal, Department of Computer Engineering, Punjabi University, India
Prof R. Naidoo, Dept of Mathematics/Center for Advanced Computer Modelling, Durban University of Technology, Durban,South Africa
Prof. Mydhili K Nair, M S Ramaiah Institute of Technology(M.S.R.I.T), Affiliated to Visweswaraiah Technological University, Bangalore, India
M. Prabu, Adhiyamaan College of Engineering/Anna University, India
Mr. Swakkhar Shatabda, Department of Computer Science and Engineering, United International University, Bangladesh
Dr. Abdur Rashid Khan, ICIT, Gomal University, Dera Ismail Khan, Pakistan
Mr. H. Abdul Shabeer, I-Nautix Technologies,Chennai, India
Dr. M. Aramudhan, Perunthalaivar Kamarajar Institute of Engineering and Technology, India
Dr. M. P. Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), India
Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran
Mr. Zeashan Hameed Khan, : Université de Grenoble, France
Prof. Anil K Ahlawat, Ajay Kumar Garg Engineering College, Ghaziabad, UP Technical University, Lucknow
Mr. Longe Olumide Babatope, University Of Ibadan, Nigeria
Associate Prof. Raman Maini, University College of Engineering, Punjabi University, India

Dr. Maslin Masrom, University Technology Malaysia, Malaysia
Sudipta Chattopadhyay, Jadavpur University, Kolkata, India
Dr. Dang Tuan NGUYEN, University of Information Technology, Vietnam National University - Ho Chi Minh City
Dr. Mary Lourde R., BITS-PILANI Dubai , UAE
Dr. Abdul Aziz, University of Central Punjab, Pakistan
Mr. Karan Singh, Gautam Budtha University, India
Mr. Avinash Pokhriyal, Uttar Pradesh Technical University, Lucknow, India
Associate Prof Dr Zuraini Ismail, University Technology Malaysia, Malaysia
Assistant Prof. Yasser M. Alginahi, College of Computer Science and Engineering, Taibah University, Madinah Munawwarah, KSA
Mr. Dakshina Ranjan Kisku, West Bengal University of Technology, India
Mr. Raman Kumar, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India
Associate Prof. Samir B. Patel, Institute of Technology, Nirma University, India
Dr. M.Munir Ahamed Rabbani, B. S. Abdur Rahman University, India
Asst. Prof. Koushik Majumder, West Bengal University of Technology, India
Dr. Alex Pappachen James, Queensland Micro-nanotechnology center, Griffith University, Australia
Assistant Prof. S. Hariharan, B.S. Abdur Rahman University, India
Asst Prof. Jasmine. K. S, R.V.College of Engineering, India
Mr Naushad Ali Mamode Khan, Ministry of Education and Human Resources, Mauritius
Prof. Mahesh Goyani, G H Patel Collge of Engg. & Tech, V.V.N, Anand, Gujarat, India
Dr. Mana Mohammed, University of Tlemcen, Algeria
Prof. Jatinder Singh, Universal Institutiion of Engg. & Tech. CHD, India
Mrs. M. Anandhavalli Gauthaman, Sikkim Manipal Institute of Technology, Majitar, East Sikkim
Dr. Bin Guo, Institute Telecom SudParis, France
Mrs. Maleika Mehr Nigar Mohamed Heenaye-Mamode Khan, University of Mauritius
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Mr. V. Bala Dhandayuthapani, Mekelle University, Ethiopia
Dr. Irfan Syamsuddin, State Polytechnic of Ujung Pandang, Indonesia
Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius
Mr. Ravi Chandiran, Zagro Singapore Pte Ltd. Singapore
Mr. Milindkumar V. Sarode, Jawaharlal Darda Institute of Engineering and Technology, India
Dr. Shamimul Qamar, KSJ Institute of Engineering & Technology, India
Dr. C. Arun, Anna University, India
Assist. Prof. M.N.Birje, Basaveshwar Engineering College, India
Prof. Hamid Reza Naji, Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran
Assist. Prof. Debasis Giri, Department of Computer Science and Engineering, Haldia Institute of Technology
Subhabrata Barman, Haldia Institute of Technology, West Bengal
Mr. M. I. Lali, COMSATS Institute of Information Technology, Islamabad, Pakistan
Dr. Feroz Khan, Central Institute of Medicinal and Aromatic Plants, Lucknow, India
Mr. R. Nagendran, Institute of Technology, Coimbatore, Tamilnadu, India
Mr. Amnach Khawne, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand

Dr. P. Chakrabarti, Sir Padampat Singhanian University, Udaipur, India
Mr. Nafiz Imtiaz Bin Hamid, Islamic University of Technology (IUT), Bangladesh.
Shahab-A. Shamshirband, Islamic Azad University, Chalous, Iran
Prof. B. Priestly Shan, Anna Univeristy, Tamilnadu, India
Venkatramreddy Velma, Dept. of Bioinformatics, University of Mississippi Medical Center, Jackson MS USA
Akshi Kumar, Dept. of Computer Engineering, Delhi Technological University, India
Dr. Umesh Kumar Singh, Vikram University, Ujjain, India
Mr. Serguei A. Mokhov, Concordia University, Canada
Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia
Dr. Awadhesh Kumar Sharma, Madan Mohan Malviya Engineering College, India
Mr. Syed R. Rizvi, Analytical Services & Materials, Inc., USA
Dr. S. Karthik, SNS College of Technology, India
Mr. Syed Qasim Bukhari, CIMET (Universidad de Granada), Spain
Mr. A.D.Potgantwar, Pune University, India
Dr. Himanshu Aggarwal, Punjabi University, India
Mr. Rajesh Ramachandran, Naipunya Institute of Management and Information Technology, India
Dr. K.L. Shunmuganathan, R.M.K Engg College, Kavaraipeitai, Chennai
Dr. Prasant Kumar Pattnaik, KIST, India.
Dr. Ch. Aswani Kumar, VIT University, India
Mr. Ijaz Ali Shoukat, King Saud University, Riyadh KSA
Mr. Arun Kumar, Sir Padam Pat Singhanian University, Udaipur, Rajasthan
Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia
Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA
Mr. Mohd Zaki Bin Mas'ud, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
Prof. Dr. R. Geetharamani, Dept. of Computer Science and Eng., Rajalakshmi Engineering College, India
Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India
Dr. S. Abdul Khader Jilani, University of Tabuk, Tabuk, Saudi Arabia
Mr. Syed Jamal Haider Zaidi, Bahria University, Pakistan
Dr. N. Devarajan, Government College of Technology, Coimbatore, Tamilnadu, INDIA
Mr. R. Jagadeesh Kannan, RMK Engineering College, India
Mr. Deo Prakash, Shri Mata Vaishno Devi University, India
Mr. Mohammad Abu Naser, Dept. of EEE, IUT, Gazipur, Bangladesh
Assist. Prof. Prasun Ghosal, Bengal Engineering and Science University, India
Mr. Md. Golam Kaosar, School of Engineering and Science, Victoria University, Melbourne City, Australia
Mr. R. Mahammad Shafi, Madanapalle Institute of Technology & Science, India
Dr. F.Sagayaraj Francis, Pondicherry Engineering College, India
Dr. Ajay Goel, HIET, Kaithal, India
Mr. Nayak Sunil Kashibarao, Bahirji Smarak Mahavidyalaya, India
Mr. Suhas J Manangi, Microsoft India
Dr. Kalyankar N. V., Yeshwant Mahavidyalaya, Nanded, India
Dr. K.D. Verma, S.V. College of Post graduate studies & Research, India
Dr. Amjad Rehman, University Technology Malaysia, Malaysia

Mr. Rachit Garg, L K College, Jalandhar, Punjab

Mr. J. William, M.A.M college of Engineering, Trichy, Tamilnadu, India

Prof. Jue-Sam Chou, Nanhua University, College of Science and Technology, Taiwan

Dr. Thorat S.B., Institute of Technology and Management, India

Mr. Ajay Prasad, Sir Padampat Singhania University, Udaipur, India

Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology & Science, India

Mr. Syed Rafiul Hussain, Ahsanullah University of Science and Technology, Bangladesh

Mrs Fazeela Tunnisa, Najran University, Kingdom of Saudi Arabia

Mrs Kavita Taneja, Maharishi Markandeshwar University, Haryana, India

Mr. Maniyar Shiraz Ahmed, Najran University, Najran, KSA

Mr. Anand Kumar, AMC Engineering College, Bangalore

Dr. Rakesh Chandra Gangwar, Beant College of Engg. & Tech., Gurdaspur (Punjab) India

Dr. V V Rama Prasad, Sree Vidyanikethan Engineering College, India

Assist. Prof. Neetesh Kumar Gupta, Technocrats Institute of Technology, Bhopal (M.P.), India

Mr. Ashish Seth, Uttar Pradesh Technical University, Lucknow, UP India

Dr. V V S S S Balaram, Sreenidhi Institute of Science and Technology, India

Mr Rahul Bhatia, Lingaya's Institute of Management and Technology, India

Prof. Niranjana Reddy, P, KITS, Warangal, India

Prof. Rakesh. Lingappa, Vijetha Institute of Technology, Bangalore, India

Dr. Mohammed Ali Hussain, Nimra College of Engineering & Technology, Vijayawada, A.P., India

Dr. A.Srinivasan, MNM Jain Engineering College, Rajiv Gandhi Salai, Thorapakkam, Chennai

Mr. Rakesh Kumar, M.M. University, Mullana, Ambala, India

Dr. Lena Khaled, Zarqa Private University, Aman, Jordan

Ms. Supriya Kapoor, Patni/Lingaya's Institute of Management and Tech., India

Dr. Tossapon Boongoen, Aberystwyth University, UK

Dr. Bilal Alatas, Firat University, Turkey

Assist. Prof. Jyoti Praaksh Singh, Academy of Technology, India

Dr. Ritu Soni, GNG College, India

Dr. Mahendra Kumar, Sagar Institute of Research & Technology, Bhopal, India.

Dr. Binod Kumar, Lakshmi Narayan College of Tech.(LNCT) Bhopal India

Dr. Muzhir Shaban Al-Ani, Amman Arab University Amman – Jordan

Dr. T.C. Manjunath, ATRIA Institute of Tech, India

Mr. Muhammad Zakarya, COMSATS Institute of Information Technology (CIIT), Pakistan

Assist. Prof. Harmunish Taneja, M. M. University, India

Dr. Chitra Dhawale, SICSR, Model Colony, Pune, India

Mrs Sankari Muthukaruppan, Nehru Institute of Engineering and Technology, Anna University, India

Mr. Aaqif Afzaal Abbasi, National University Of Sciences And Technology, Islamabad

Prof. Ashutosh Kumar Dubey, Trinity Institute of Technology and Research Bhopal, India

Mr. G. Appasami, Dr. Pauls Engineering College, India

Mr. M Yasin, National University of Science and Tech, Karachi (NUST), Pakistan

Mr. Yaser Miaji, University Utara Malaysia, Malaysia

Mr. Shah Ahsanul Haque, International Islamic University Chittagong (IIUC), Bangladesh

Prof. (Dr) Syed Abdul Sattar, Royal Institute of Technology & Science, India
Dr. S. Sasikumar, Roever Engineering College
Assist. Prof. Monit Kapoor, Maharishi Markandeshwar University, India
Mr. Nwaocha Vivian O, National Open University of Nigeria
Dr. M. S. Vijaya, GR Govindarajulu School of Applied Computer Technology, India
Assist. Prof. Chakresh Kumar, Manav Rachna International University, India
Mr. Kunal Chadha , R&D Software Engineer, Gemalto, Singapore
Mr. Mueen Uddin, Universiti Teknologi Malaysia, UTM , Malaysia
Dr. Dhuha Basheer abdullah, Mosul university, Iraq
Mr. S. Audithan, Annamalai University, India
Prof. Vijay K Chaudhari, Technocrats Institute of Technology , India
Associate Prof. Mohd Ilyas Khan, Technocrats Institute of Technology , India
Dr. Vu Thanh Nguyen, University of Information Technology, HoChiMinh City, VietNam
Assist. Prof. Anand Sharma, MITS, Lakshmangarh, Sikar, Rajasthan, India
Prof. T V Narayana Rao, HITAM Engineering college, Hyderabad
Mr. Deepak Gour, Sir Padampat Singhania University, India
Assist. Prof. Amutharaj Joyson, Kalasalingam University, India
Mr. Ali Balador, Islamic Azad University, Iran
Mr. Mohit Jain, Maharaja Surajmal Institute of Technology, India
Mr. Dilip Kumar Sharma, GLA Institute of Technology & Management, India
Dr. Debojyoti Mitra, Sir padampat Singhania University, India
Dr. Ali Dehghantanha, Asia-Pacific University College of Technology and Innovation, Malaysia
Mr. Zhao Zhang, City University of Hong Kong, China
Prof. S.P. Setty, A.U. College of Engineering, India
Prof. Patel Rakeshkumar Kantilal, Sankalchand Patel College of Engineering, India
Mr. Biswajit Bhowmik, Bengal College of Engineering & Technology, India
Mr. Manoj Gupta, Apex Institute of Engineering & Technology, India
Assist. Prof. Ajay Sharma, Raj Kumar Goel Institute Of Technology, India
Assist. Prof. Ramveer Singh, Raj Kumar Goel Institute of Technology, India
Dr. Hanan Elazhary, Electronics Research Institute, Egypt
Dr. Hosam I. Faiq, USM, Malaysia
Prof. Dipti D. Patil, MAEER's MIT College of Engg. & Tech, Pune, India
Assist. Prof. Devendra Chack, BCT Kumaon engineering College Dwarahat Almora, India
Prof. Manpreet Singh, M. M. Engg. College, M. M. University, India
Assist. Prof. M. Sadiq ali Khan, University of Karachi, Pakistan
Mr. Prasad S. Halgaonkar, MIT - College of Engineering, Pune, India
Dr. Imran Ghani, Universiti Teknologi Malaysia, Malaysia
Prof. Varun Kumar Kakar, Kumaon Engineering College, Dwarahat, India
Assist. Prof. Nisheeth Joshi, Apaji Institute, Banasthali University, Rajasthan, India
Associate Prof. Kunwar S. Vaisla, VCT Kumaon Engineering College, India
Prof Anupam Choudhary, Bhilai School Of Engg.,Bhilai (C.G.),India
Mr. Divya Prakash Shrivastava, Al Jabal Al garbi University, Zawya, Libya

Associate Prof. Dr. V. Radha, Avinashilingam Deemed university for women, Coimbatore.
Dr. Kasarapu Ramani, JNT University, Anantapur, India
Dr. Anuraag Awasthi, Jayoti Vidyapeeth Womens University, India
Dr. C G Ravichandran, R V S College of Engineering and Technology, India
Dr. Mohamed A. Deriche, King Fahd University of Petroleum and Minerals, Saudi Arabia
Mr. Abbas Karimi, Universiti Putra Malaysia, Malaysia
Mr. Amit Kumar, Jaypee University of Engg. and Tech., India
Dr. Nikolai Stoianov, Defense Institute, Bulgaria
Assist. Prof. S. Ranichandra, KSR College of Arts and Science, Tiruchencode
Mr. T.K.P. Rajagopal, Diamond Horse International Pvt Ltd, India
Dr. Md. Ekramul Hamid, Rajshahi University, Bangladesh
Mr. Hemanta Kumar Kalita , TATA Consultancy Services (TCS), India
Dr. Messaouda Azzouzi, Ziane Achour University of Djelfa, Algeria
Prof. (Dr.) Juan Jose Martinez Castillo, "Gran Mariscal de Ayacucho" University and Acantelys research Group, Venezuela
Dr. Jatinderkumar R. Saini, Narmada College of Computer Application, India
Dr. Babak Bashari Rad, University Technology of Malaysia, Malaysia
Dr. Nighat Mir, Effat University, Saudi Arabia
Prof. (Dr.) G.M.Nasira, Sasurie College of Engineering, India
Mr. Varun Mittal, Gemalto Pte Ltd, Singapore
Assist. Prof. Mrs P. Banumathi, Kathir College Of Engineering, Coimbatore
Assist. Prof. Quan Yuan, University of Wisconsin-Stevens Point, US
Dr. Pranam Paul, Narula Institute of Technology, Agarpara, West Bengal, India
Assist. Prof. J. Ramkumar, V.L.B Janakiammal college of Arts & Science, India
Mr. P. Sivakumar, Anna university, Chennai, India
Mr. Md. Humayun Kabir Biswas, King Khalid University, Kingdom of Saudi Arabia
Mr. Mayank Singh, J.P. Institute of Engg & Technology, Meerut, India
HJ. Kamaruzaman Jusoff, Universiti Putra Malaysia
Mr. Nikhil Patrick Lobo, CADES, India
Dr. Amit Wason, Rayat-Bahra Institute of Engineering & Boi-Technology, India
Dr. Rajesh Shrivastava, Govt. Benazir Science & Commerce College, Bhopal, India
Assist. Prof. Vishal Bharti, DCE, Gurgaon
Mrs. Sunita Bansal, Birla Institute of Technology & Science, India
Dr. R. Sudhakar, Dr.Mahalingam college of Engineering and Technology, India
Dr. Amit Kumar Garg, Shri Mata Vaishno Devi University, Katra(J&K), India
Assist. Prof. Raj Gaurang Tiwari, AZAD Institute of Engineering and Technology, India
Mr. Hamed Taherdoost, Tehran, Iran
Mr. Amin Daneshmand Malayeri, YRC, IAU, Malayer Branch, Iran
Mr. Shantanu Pal, University of Calcutta, India
Dr. Terry H. Walcott, E-Promag Consultancy Group, United Kingdom
Dr. Ezekiel U OKIKE, University of Ibadan, Nigeria
Mr. P. Mahalingam, Caledonian College of Engineering, Oman

Dr. Mahmoud M. A. Abd Ellatif, Mansoura University, Egypt
Prof. Kunwar S. Vaisla, BCT Kumaon Engineering College, India
Prof. Mahesh H. Panchal, Kalol Institute of Technology & Research Centre, India
Mr. Muhammad Asad, Technical University of Munich, Germany
Mr. AliReza Shams Shafigh, Azad Islamic university, Iran
Prof. S. V. Nagaraj, RMK Engineering College, India
Mr. Ashikali M Hasan, Senior Researcher, CelNet security, India
Dr. Adnan Shahid Khan, University Technology Malaysia, Malaysia
Mr. Prakash Gajanan Burade, Nagpur University/ITM college of engg, Nagpur, India
Dr. Jagdish B. Helonde, Nagpur University/ITM college of engg, Nagpur, India
Professor, Doctor BOUHORMA Mohammed, Univertsity Abdelmalek Essaadi, Morocco
Mr. K. Thirumalaivasan, Pondicherry Engg. College, India
Mr. Umbarkar Anantkumar Janardan, Walchand College of Engineering, India
Mr. Ashish Chaurasia, Gyan Ganga Institute of Technology & Sciences, India
Mr. Sunil Taneja, Kurukshetra University, India
Mr. Fauzi Adi Rafrastara, Dian Nuswantoro University, Indonesia
Dr. Yaduvir Singh, Thapar University, India
Dr. Ioannis V. Koskosas, University of Western Macedonia, Greece
Dr. Vasantha Kalyani David, Avinashilingam University for women, Coimbatore
Dr. Ahmed Mansour Manasrah, Universiti Sains Malaysia, Malaysia
Miss. Nazanin Sadat Kazazi, University Technology Malaysia, Malaysia
Mr. Saeed Rasouli Heikalabad, Islamic Azad University - Tabriz Branch, Iran
Assoc. Prof. Dharendra Mishra, SVKM's NMIMS University, India
Prof. Shapoor Zarei, UAE Inventors Association, UAE
Prof. B.Raja Sarath Kumar, Lenora College of Engineering, India
Dr. Bashir Alam, Jamia millia Islamia, Delhi, India
Prof. Anant J Umbarkar, Walchand College of Engg., India
Assist. Prof. B. Bharathi, Sathyabama University, India
Dr. Fokrul Alom Mazarbhuiya, King Khalid University, Saudi Arabia
Prof. T.S.Jeyali Laseeth, Anna University of Technology, Tirunelveli, India
Dr. M. Balraju, Jawahar Lal Nehru Technological University Hyderabad, India
Dr. Vijayalakshmi M. N., R.V.College of Engineering, Bangalore
Prof. Walid Moudani, Lebanese University, Lebanon
Dr. Saurabh Pal, VBS Purvanchal University, Jaunpur, India
Associate Prof. Suneet Chaudhary, Dehradun Institute of Technology, India
Associate Prof. Dr. Manuj Darbari, BBD University, India
Ms. Prema Selvaraj, K.S.R College of Arts and Science, India
Assist. Prof. Ms.S.Sasikala, KSR College of Arts & Science, India
Mr. Sukhvinder Singh Deora, NC Institute of Computer Sciences, India
Dr. Abhay Bansal, Amity School of Engineering & Technology, India
Ms. Sumita Mishra, Amity School of Engineering and Technology, India
Professor S. Viswanadha Raju, JNT University Hyderabad, India

Mr. Asghar Shahrzad Khashandarag, Islamic Azad University Tabriz Branch, India
Mr. Manoj Sharma, Panipat Institute of Engg. & Technology, India
Mr. Shakeel Ahmed, King Faisal University, Saudi Arabia
Dr. Mohamed Ali Mahjoub, Institute of Engineer of Monastir, Tunisia
Mr. Adri Jovin J.J., SriGuru Institute of Technology, India
Dr. Sukumar Senthilkumar, Universiti Sains Malaysia, Malaysia
Mr. Rakesh Bharati, Dehradun Institute of Technology Dehradun, India
Mr. Shervan Fekri Ershad, Shiraz International University, Iran
Mr. Md. Safiqul Islam, Daffodil International University, Bangladesh
Mr. Mahmudul Hasan, Daffodil International University, Bangladesh
Prof. Mandakini Tayade, UIT, RGTU, Bhopal, India
Ms. Sarla More, UIT, RGTU, Bhopal, India
Mr. Tushar Hrishikesh Jaware, R.C. Patel Institute of Technology, Shirpur, India
Ms. C. Divya, Dr G R Damodaran College of Science, Coimbatore, India
Mr. Fahimuddin Shaik, Annamacharya Institute of Technology & Sciences, India
Dr. M. N. Giri Prasad, JNTUCE,Pulivendula, A.P., India
Assist. Prof. Chintan M Bhatt, Charotar University of Science And Technology, India
Prof. Sahista Machchhar, Marwadi Education Foundation's Group of institutions, India
Assist. Prof. Navnish Goel, S. D. College Of Enginnering & Technology, India
Mr. Khaja Kamaluddin, Sirt University, Sirt, Libya
Mr. Mohammad Zaidul Karim, Daffodil International, Bangladesh
Mr. M. Vijayakumar, KSR College of Engineering, Tiruchengode, India
Mr. S. A. Ahsan Rajon, Khulna University, Bangladesh
Dr. Muhammad Mohsin Nazir, LCW University Lahore, Pakistan
Mr. Mohammad Asadul Hoque, University of Alabama, USA
Mr. P.V.Sarathchand, Indur Institute of Engineering and Technology, India
Mr. Durgesh Samadhiya, Chung Hua University, Taiwan
Dr Venu Kuthadi, University of Johannesburg, Johannesburg, RSA
Dr. (Er) Jasvir Singh, Guru Nanak Dev University, Amritsar, Punjab, India
Mr. Jasmin Cosic, Min. of the Interior of Una-sana canton, B&H, Bosnia and Herzegovina
Dr S. Rajalakshmi, Botho College, South Africa
Dr. Mohamed Sarrab, De Montfort University, UK
Mr. Basappa B. Kodada, Canara Engineering College, India
Assist. Prof. K. Ramana, Annamacharya Institute of Technology and Sciences, India
Dr. Ashu Gupta, Apeejay Institute of Management, Jalandhar, India
Assist. Prof. Shaik Rasool, Shadan College of Engineering & Technology, India
Assist. Prof. K. Suresh, Annamacharya Institute of Tech & Sci. Rajampet, AP, India
Dr . G. Singaravel, K.S.R. College of Engineering, India
Dr B. G. Geetha, K.S.R. College of Engineering, India
Assist. Prof. Kavita Choudhary, ITM University, Gurgaon
Dr. Mehrdad Jalali, Azad University, Mashhad, Iran
Megha Goel, Shamli Institute of Engineering and Technology, Shamli, India

Mr. Chi-Hua Chen, Institute of Information Management, National Chiao-Tung University, Taiwan (R.O.C.)

Assoc. Prof. A. Rajendran, RVS College of Engineering and Technology, India

Assist. Prof. S. Jaganathan, RVS College of Engineering and Technology, India

Assoc. Prof. A S N Chakravarthy, Sri Aditya Engineering College, India

Assist. Prof. Deepshikha Patel, Technocrat Institute of Technology, India

Assist. Prof. Maram Balajee, GMRIT, India

Assist. Prof. Monika Bhatnagar, TIT, India

Prof. Gaurang Panchal, Charotar University of Science & Technology, India

Prof. Anand K. Tripathi, Computer Society of India

Prof. Jyoti Chaudhary, High Performance Computing Research Lab, India

Assist. Prof. Supriya Raheja, ITM University, India

Dr. Pankaj Gupta, Microsoft Corporation, U.S.A.

Assist. Prof. Panchamukesh Chandaka, Hyderabad Institute of Tech. & Management, India

Prof. Mohan H.S, SJB Institute Of Technology, India

Mr. Hossein Malekinezhad, Islamic Azad University, Iran

Mr. Zatin Gupta, Universti Malaysia, Malaysia

Assist. Prof. Amit Chauhan, Phonics Group of Institutions, India

Assist. Prof. Ajal A. J., METS School Of Engineering, India

Mrs. Omowunmi Omobola Adeyemo, University of Ibadan, Nigeria

Dr. Bharat Bhushan Agarwal, I.F.T.M. University, India

Md. Nazrul Islam, University of Western Ontario, Canada

Tushar Kanti, L.N.C.T, Bhopal, India

Er. Aumreesh Kumar Saxena, SIRTs College Bhopal, India

Mr. Mohammad Monirul Islam, Daffodil International University, Bangladesh

Dr. Kashif Nisar, University Utara Malaysia, Malaysia

Dr. Wei Zheng, Rutgers Univ/ A10 Networks, USA

Associate Prof. Rituraj Jain, Vyas Institute of Engg & Tech, Jodhpur – Rajasthan

Assist. Prof. Apoorvi Sood, I.T.M. University, India

Dr. Kayhan Zrar Ghafoor, University Technology Malaysia, Malaysia

Mr. Swapnil Sonar, Truba Institute College of Engineering & Technology, Indore, India

Ms. Yogita Gigras, I.T.M. University, India

Associate Prof. Neelima Sadineni, Pydha Engineering College, India Pydha Engineering College

Assist. Prof. K. Deepika Rani, HITAM, Hyderabad

Ms. Shikha Maheshwari, Jaipur Engineering College & Research Centre, India

Prof. Dr V S Giridhar Akula, Avanthi's Scientific Tech. & Research Academy, Hyderabad

Prof. Dr.S.Saravanan, Muthayammal Engineering College, India

Mr. Mehdi Golsorkhatabar Amiri, Islamic Azad University, Iran

Prof. Amit Sadanand Savyanavar, MITCOE, Pune, India

Assist. Prof. P.Oliver Jayaprakash, Anna University, Chennai

Assist. Prof. Ms. Sujata, ITM University, Gurgaon, India

Dr. Asoke Nath, St. Xavier's College, India

Mr. Masoud Rafighi, Islamic Azad University, Iran

Assist. Prof. RamBabu Pemula, NIMRA College of Engineering & Technology, India
Assist. Prof. Ms Rita Chhikara, ITM University, Gurgaon, India
Mr. Sandeep Maan, Government Post Graduate College, India
Prof. Dr. S. Muralidharan, Mepco Schlenk Engineering College, India
Associate Prof. T.V.Sai Krishna, QIS College of Engineering and Technology, India
Mr. R. Balu, Bharathiar University, Coimbatore, India
Assist. Prof. Shekhar. R, Dr.SM College of Engineering, India
Prof. P. Senthilkumar, Vivekanandha Institute of Engineering And Technology For Woman, India
Mr. M. Kamarajan, PSNA College of Engineering & Technology, India
Dr. Angajala Srinivasa Rao, Jawaharlal Nehru Technical University, India
Assist. Prof. C. Venkatesh, A.I.T.S, Rajampet, India
Mr. Afshin Rezakhani Roozbahani, Ayatollah Boroujerdi University, Iran
Mr. Laxmi chand, SCTL, Noida, India
Dr. Dr. Abdul Hannan, Vivekanand College, Aurangabad
Prof. Mahesh Panchal, KITRC, Gujarat
Dr. A. Subramani, K.S.R. College of Engineering, Tiruchengode
Assist. Prof. Prakash M, Rajalakshmi Engineering College, Chennai, India
Assist. Prof. Akhilesh K Sharma, Sir Padampat Singhanian University, India
Ms. Varsha Sahni, Guru Nanak Dev Engineering College, Ludhiana, India
Associate Prof. Trilochan Rout, NM Institute Of Engineering And Technology, India
Mr. Srikantha Kumar Mohapatra, NMIET, Orissa, India
Mr. Waqas Haider Bangyal, Iqra University Islamabad, Pakistan
Dr. S. Vijayaragavan, Christ College of Engineering and Technology, Pondicherry, India
Prof. Elboukhari Mohamed, University Mohammed First, Oujda, Morocco
Dr. Muhammad Asif Khan, King Faisal University, Saudi Arabia
Dr. Nagy Ramadan Darwish Omran, Cairo University, Egypt.
Assistant Prof. Anand Nayyar, KCL Institute of Management and Technology, India
Mr. G. Premsankar, Ericsson, India
Assist. Prof. T. Hemalatha, VELS University, India
Prof. Tejaswini Apte, University of Pune, India
Dr. Edmund Ng Giap Weng, Universiti Malaysia Sarawak, Malaysia
Mr. Mahdi Nouri, Iran University of Science and Technology, Iran
Associate Prof. S. Asif Hussain, Annamacharya Institute of technology & Sciences, India
Mrs. Kavita Pabreja, Maharaja Surajmal Institute (an affiliate of GGSIP University), India
Mr. Vorugunti Chandra Sekhar, DA-IICT, India
Mr. Muhammad Najmi Ahmad Zabidi, Universiti Teknologi Malaysia, Malaysia
Dr. Aderemi A. Atayero, Covenant University, Nigeria
Assist. Prof. Osama Sohaib, Balochistan University of Information Technology, Pakistan
Assist. Prof. K. Suresh, Annamacharya Institute of Technology and Sciences, India
Mr. Hassen Mohammed Abdullah Alsafi, International Islamic University Malaysia (IIUM) Malaysia
Mr. Robail Yasrab, Virtual University of Pakistan, Pakistan
Mr. R. Balu, Bharathiar University, Coimbatore, India

Prof. Anand Nayyar, KCL Institute of Management and Technology, Jalandhar
Assoc. Prof. Vivek S Deshpande, MIT College of Engineering, India
Prof. K. Saravanan, Anna university Coimbatore, India
Dr. Ravendra Singh, MJP Rohilkhand University, Bareilly, India
Mr. V. Mathivanan, IBRA College of Technology, Sultanate of OMAN
Assoc. Prof. S. Asif Hussain, AITS, India
Assist. Prof. C. Venkatesh, AITS, India
Mr. Sami Ulhaq, SZABIST Islamabad, Pakistan
Dr. B. Justus Rabi, Institute of Science & Technology, India
Mr. Anuj Kumar Yadav, Dehradun Institute of technology, India
Mr. Alejandro Mosquera, University of Alicante, Spain
Assist. Prof. Arjun Singh, Sir Padampat Singhanian University (SPSU), Udaipur, India
Dr. Smriti Agrawal, JB Institute of Engineering and Technology, Hyderabad
Assist. Prof. Swathi Sambangi, Visakha Institute of Engineering and Technology, India
Ms. Prabhjot Kaur, Guru Gobind Singh Indraprastha University, India
Mrs. Samaher AL-Hothali, Yanbu University College, Saudi Arabia
Prof. Rajneeshkaur Bedi, MIT College of Engineering, Pune, India
Mr. Hassen Mohammed Abdullah Alsafi, International Islamic University Malaysia (IIUM)
Dr. Wei Zhang, Amazon.com, Seattle, WA, USA
Mr. B. Santhosh Kumar, C S I College of Engineering, Tamil Nadu
Dr. K. Reji Kumar, , N S S College, Pandalam, India
Assoc. Prof. K. Seshadri Sastry, EIILM University, India
Mr. Kai Pan, UNC Charlotte, USA
Mr. Ruikar Sachin, SGGSIET, India
Prof. (Dr.) Vinodani Katiyar, Sri Ramswaroop Memorial University, India
Assoc. Prof., M. Giri, Sreenivasa Institute of Technology and Management Studies, India
Assoc. Prof. Labib Francis Gergis, Misr Academy for Engineering and Technology (MET), Egypt
Assist. Prof. Amanpreet Kaur, ITM University, India
Assist. Prof. Anand Singh Rajawat, Shri Vaishnav Institute of Technology & Science, Indore
Mrs. Hadeel Saleh Haj Aliwi, Universiti Sains Malaysia (USM), Malaysia
Dr. Abhay Bansal, Amity University, India
Dr. Mohammad A. Mezher, Fahad Bin Sultan University, KSA
Assist. Prof. Nidhi Arora, M.C.A. Institute, India
Prof. Dr. P. Suresh, Karpagam College of Engineering, Coimbatore, India

CALL FOR PAPERS
International Journal of Computer Science and Information Security
January - December
IJCSIS 2012
ISSN: 1947-5500
<http://sites.google.com/site/ijcsis/>

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc.), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on

its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embedded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at <http://sites.google.com/site/ijcsis/authors-notes> .



© IJCSIS PUBLICATION 2012
ISSN 1947 5500